Lezione 9 - Errore di Interpolazione e Minimi Quadrati

Errore di interpolazione

Sia $I = [x_0, x_n]$, e sia $f \in C^0([x_0, x_n])$.
Considerando $\{(x_i, \underbrace{y_i}_{f(x_i)})\}_{i=0}^n$ con x_i distanti.

Se $f \in C^{n+1}(\bar{I})$ allora $\forall x \in I \; \exists \alpha = \alpha(x) \in I$ tale che:

$$Enf(x) = f(x) - \Pi_n f(x) = \frac{f^{n+1}(\alpha(x))}{(n+1)!} \cdot \prod_{i=0}^{n}(x - x_i)$$

L'errore è un valore preciso, non è una disuguaglianza. L'unico problema è che non sappiamo cosa sia α e la funzione deve esser n+1 volte regolare che è un molto restrittivo sulla funzioni a cui possiamo applicarlo.

L'unico cosa che sappiamo per fatto è che a tutti i nodi l'errore e nullo per come abbiamo costruito la funzione interpolante.

Per rendere questo errore più concreto, cerchiamo il massimo di questo errore:

$$\underset{x \in I}{max}|Enf(x)| \leq \underset{x \in I}{max}\frac{|f^{n+1}(x)|}{(n+1)!} \cdot |\prod_{i=0}^{n}(x - x_i)|$$

Nel momento in cui lo massimizziamo, l'α si rimuove, questo perché o alfa(x) stesso e il punto di massimo, o un'altro x lo è, quindi ponendo x come quel punto dove lo è l'errore alla destra potrà solo esser maggiore o uguale all'errore ad $\alpha(x)$ che è il massimo errore della interpolazione. (Non ho spiegato bene)

Nodi Equidistanti

Prendendo nodi equidistanti riusciamo a semplificare ulteriormente il calcolo dell'errore dell'interpolazione.

Prendiamo la distanza tra due nodi come:

$$h = \frac{x_n - x_i}{n}$$

Per generalizzare avremmo che:

$$x_k = x_{k-1} + h \rightarrow \text{per } k = 1, \ldots, n$$

O possiamo scriverlo come:

$$x_k = x_0 + kh \rightarrow \text{con } k = 0, \ldots, n$$

Se i nodi sono equispaziati, si può dimostrare che:

$$|\prod_{i=0}^{n}(x-x)i| \leq \frac{h^{n+1}}{4}n!$$

$$\implies \max_{x\in I}|Enf(x)| \leq \max_{x\in I}\frac{|f^{n+1}(x)|}{\cancel{(n+1)!}} \cdot \frac{h^{n+1}}{4}\cancel{n!}$$

$$\leq \max_{x\in I}\frac{|f^{n+1}|}{n+1} \cdot \frac{h^{n+1}}{4}$$

Per i nodi equispaziati, possiamo dire che:

$$\max_{x\in I}|Enf(x)| \leq \frac{h^{n+1}}{4(n+1)} \cdot \max_{x\in I}|f^{n+1}(x)|$$

Più nodi abbiamo, meglio si approssima la funzione.
Per $n \to \infty \cancel{\implies} Enf \to 0$, questo dipende dall'ordine delle due parti della funzione.

Il blocco A:

$$\frac{h^{n+1}}{4(n+1)} \to 0 \text{ per } n \to \infty$$

Invece il blocco B:

$$\max_{x\in I}|f^{n+1}(x)|$$

Ha 3 possibilità:

- può andare a 0
- può andare a una costante
- può andare a ∞

I primi due casi non sono problematici, il terzo lo può esser.

Nel caso in cui il blocco B tenda a ∞, la determinazione della convergenza dell'errore dipende dall'ordine di convergenza del blocco A e B.

Se il blocco a destra ha ordine maggiore, ordine va male, come si manifesta questo è una oscillazione che peggiora più ci si avvicina agli estremi della funzione.

Un esempio sarebbe:

Questo fenomeno di oscillazione a causa dell'errore è noto come il fenomeno di Runge, e ci sono diversi modi per correggerlo.

Metodi per correre l'errore

Ci son diversi modi per diminuire l'effetto di Runge. Principalmente questi metodi operano sulla distribuzione di nodi nell'intervallo e considerando il l'andamento della funzione stessa per scegliere i nodi.

Guardiamo 3 metodi:

- Nodi di Chebyshev
- Nodi adattativi

- Interpolazione a tratti

1. Nodi di Chebyshev-Gauss-Lobatto

Con i nodi di Chevyshev aumentiamo la fittezza dei nodi agli estremi, più nodi abbiamo agli estremi più forziamo la adiacenza alla funzione vera.

Nell'intervallo $[-1, 1]$, per $i = 0, \ldots, n$, prendiamo i punti:

$$x_i^* = -\cos\left(\frac{\pi i}{n}\right)$$

I punti risultano esser più vicini agli estremi e più lontani al centro.

I punti di Chebyshev usano questa distribuzione nell'intervallo intero che stiamo guardando, facendo:

$$x_i^C = \frac{b+a}{2} + \frac{b-a}{2} \cdot x_i^*$$

Con questi nodi si crea un teorema, che dice:

$$\text{Se } f \in C^1(\bar{I}) \implies \Pi_n f \to f \text{ per } n \to \infty$$

Questo significa che per un $\Pi_n f$ generato con i nodi di Chebyshev serve meno una funzione meno restrittiva per fare una interpolazione.

L'errore con Chevyshev è contenuto ma il suo problema è che funziona solo con funzioni specifiche.

2. Nodi Adattativi

I nodi adattativi sono nodi per usando la derivata della funzione per determinare la distribuzione dei nodi.

Usiamo una discretizzazione lasca dove ci sono le sezioni piatte e aumentiamo la fittezza dove il gradino è più alto.

Questo è il metodo usato dalla funzione plot in matlab.

3. Interpolazione a tratti (piecewise)

Questa è diversa dalla interpolazione di lagrange, che è globale, l'interpolazione a tratti è un insieme di interpolazioni locali.

Anziché usare tutti i nodi insieme, ne guardiamo un po' alla volta, dividendo in sotto-intervalli e poi facendo interpolazioni, sotto=intervallo per sotto-intervallo.

Se usiamo 2 nodi per intervallo, troviamo rette in ogni sotto-intervallo, visto che le rette non possono oscillare, più nodi abbiamo meno l'errore sarà.

Prendendo x_i distinti.
Il nostro intervallo sarà:

$$I_i[x_i, x_{i+1}] \text{ con } i = 0, \ldots, n-1$$

La larghezza di ogni intervallo sarà $h_i = x_{i+1} - x_i$, e definiamo $H = max_i h_i$.

Prendiamo $\Pi_1^H f \in C^0(\bar{I})$, dove 1 indica che è lineare, H significa che tiene a conto di H e che stiamo facendo una interpolazione a tratti.

Imponendo una ristrizione ad I_i, prendiamo $\Pi_1^H f|_{I_i} \in \mathbb{P}_1(\bar{I}_i)$ tale che $\Pi_1^H f(x_i) = y_i$ per $i = 0, \dots, n$

La caratterizzazione analitica di funzione di interpolazione intervallo per intervallo è:

$$\Pi_1^H f|_{I_i}(x) = f(x_i) + \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}(x - x_i)$$

Fogliamo accertarci che sia una soluzione alle onde, l'errore massimo di interpolazione è:

$$\max_{x \in I_i} |[f(x) - \Pi_1^H f(x)]|_{I_i}|$$

Globalmente i nodi non sono equispaziati ma intervallo per intervallo lo sono, quindi visto che stiamo guardando a questa scala va bene

L'errore sarà:

$$\max_{x \in I_i} |[f(x) - \Pi_1^H f(x)]|_{I_i}| \le \frac{h_i^2}{4 * 2} \cdot \max_{x \in I_i} |f''(x)|$$

Guardando l'intervallo intero, massimizziamo tutte le parti:

$$\max_{x \in I} |f(x) - \Pi_1^H f(x)| \le \frac{H^2}{8} \max_{x \in I} |f''(x)|$$

Per $n \to \infty$, H $\to$ 0, invece la seconda parte è una costante perché non è più dipendente da n come prima. Rimane la seconda derivata perché stiamo trovando il valore per ogni sotto-intervallo della seconda derivata, per quanto ora stiamo guardando l'intervallo intero, non cambia il fatto che il nostro polinomio interpolante è una serie di polinomi interpolanti di lineari che usano la seconda derivata per stimare l'errore quindi usiamo anche lungo tutto l'intervallo la seconda derivata.

Il fatto che H $\to$ 0, mentre l'altra parte rimane constante, significa che l'errore anche lui tende a 0, risolvendo il problema delle oscillazioni.

Minimi Quadrati

Nel caso di molto dati molto vicini, la interpolazione non è la scelta migliore. Cerchiamo qualcosa che definisce l'andamento medio dei dati.

Questa relazione che troviamo ci permette di estrapolare un andamento generale invece di avere valori assoluti per tutti i punti, permettendoci anche ci determinare cosa succederà se volessimo estendere i dati.

Dato un insieme di coppie di dati $\{(x_i, y_i)\}_{i=0}^n$ con x_i distinti.
Dobbiamo cercare $\tilde{f} \in \mathbb{P}_m$ di solito con $m \ll n$ tale che:

$$\sum_{i=0}^n [y_i - \tilde{f}(x_i)]^2 \le \sum_{i=0}^n \underbrace{[y_i - p_m(x_i)]}_{\forall p_m \in \mathbb{P}_m}{}^2$$

Cioè, gli scarti quadrati associati associati a $\tilde{f}$ sono minori o uguali agli scarti quadratici associati a p_m generico. $\tilde{f}$ è fra tutti i p_m quello che minimizza lo somma degli scarti quadratici.

Osservazioni

- m è qualsiasi

- se $\widetilde{f}$ esiste (non è garantito) è detto approssimazione ai minimi quadrati
- per m = 1, $\widetilde{f}$ è una retta, ed è detta la retta di regressione
- se prendiamo m = n, ogni punto è nodo, quindi la somma quadratica è 0 per ogni x_i, allora $\widetilde{f}(x_i) = y_i$, allora stiamo facendo un'interpolazione. L'interpolazione è un caso speciale dei minimi quadrati quando m = n.
- Di solito $\widetilde{f}(x_i) \neq y_i$ per $i = 0, \ldots, n$

Costruzione del Polinomio e il Problema di Minimizzazione

$$\widetilde{f}(x) = a_0 + a_1 x + \cdots + a_m x^m$$

Quello che vogliamo trovare sono i valori, $a_0, a_1, \ldots, a_m$

$$p_m(x) = b_0 + b_1 x + \cdots + b_m x^m$$

$$\sum_{i=0}^{n} [y_i - a_0 - a_1 x_i - \cdots - a_m x_i]^2 \leq \underbrace{\sum_{i=0}^{n} [y_i - b_0 - b_1 x_i^2 - \cdots - b_m x_i^m]^2}_{\phi(b_0, b_1, \ldots, b_m)}$$

Il minimo è raggiunto quando $b_0 = a_0$, $b_1 = a_1$,. . ., $b_m = a_m$

$$\implies \phi(a_o, a_1, \ldots, a_m) = \min_{b_0, b_1, \ldots, b_m} \phi(b_0, b_1, \ldots, b_m)$$

Fissando m = 1:

$$\phi(a_0, a_1) = \min_{b_0, b_1} \phi(b_0, b_1)$$

Questa è la funzione per la retta di minimizzazione.

$$\phi(b_0, b_1) = \sum_{i=0}^{n} [y_i - b_0 - b_1 x_i]^2$$

$$\phi(b_0, b_1) = \sum_{i=0}^{n} [y_i^2 + b_0^2 + b_1^2 x_i^2 - 2 y_i b_0 - 2 y_i b_1 x_i - 2 b_0 b_1 x_i]$$

Trovare a_0 e a_1 allora significa trovare il minimo di paraboloide.

$$\frac{\partial \phi}{\partial b_0}(a_0, a_1) = 0$$
$$\frac{\partial \phi}{\partial b_1}(a_0, a_1) = 0$$

$$\frac{\partial \phi}{\partial b_0}(b_0, b_1) = \sum_{i=0}^{n} [2 b_0 - 2 y_i + 2 b_1 x_i]$$
$$\frac{\partial \phi}{\partial b_1}(b_0, b_1) = \sum_{i=0}^{n} [2 b_1 x_i^2 - 2 y_i x_i + 2 b_o x_i]$$

Quando $(b_0, b_1) = (a_0, a_1)$, le ultime due funzioni si annullano. Queste sono due equazioni lineari con incognite a_0 e a_1, che possiamo trovare risolvendo il sistema.