

CAPITOLO V

APPROSSIMAZIONE DELLE FUNZIONI

1. INTRODUZIONEApprossimazione con polinomi algebrici.

Sebbene noi siamo capaci di operare comunemente con funzioni del tipo $\sin(x)$, $\cos(x)$, e^x , $\tanh(x)$, \sqrt{x} ecc, in quanto ne conosciamo molte proprietà che ci consentono di manipolarle e di metterle in relazione tra loro, la loro utilizzazione sul piano quantitativo è limitata dal fatto che esse sono definite attraverso un procedimento di passaggio al limite, in particolare con una serie convergente di potenze. Per esempio per ogni x , la funzione $\sin(x)$ può essere definita attraverso la sua serie di Mac Laurin:

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \dots$$

per definire la quale occorre conoscere i valori di $\sin(x)$ e di tutte le sue derivate in 0.

Ciò significa che, in generale, non si può esprimere il valore di $\sin(x)$ in termini finiti rispetto ad x . D'altra parte per ogni \bar{x} è possibile ottenere una approssimazione di $\sin(\bar{x})$, con errore inferiore a qualunque tolleranza prefissata a priori, attraverso un opportuno troncamento della serie calcolata in \bar{x} . Inoltre, per k sufficientemente grande, il polinomio

$$p_{2k+1}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots (-1)^k \frac{x^{2k+1}}{(2k+1)!}$$

è capace di approssimare $\sin(x)$ al di sotto della tolleranza data su tutto l'insieme di periodicità $[-\pi, \pi]$. Quest'ultimo fatto è garantito dalla convergenza uniforme dello sviluppo di Mac Laurin. Dunque il polinomio $p_{2k+1}(x)$ rappresenta un concreto "sostituto" della funzione $\sin(x)$ a meno di quella tolleranza.

Più in generale, può accadere che per una funzione $f(x)$ assegnata, si desideri disporre di una approssimazione uniforme di tipo polinomiale che risulta più semplice da tabulare, da integrare, da derivare ecc.

Se la funzione è sviluppabile in serie di Taylor, ancora una volta il polinomio :

$$p_n(x) = f(x_0) + (x - x_0)f'(x_0) + \dots + \frac{(x - x_0)^n}{n!} f^{(n)}(x_0)$$

fornisce una possibile approssimazione purchè si disponga effettivamente delle sue derivate successive nel punto x_0 . Quest'ultima condizione limita notevolmente, sul piano numerico, l'uso dei polinomi di Taylor come polinomi approssimanti. Inoltre è noto che essi forniscono delle buone approssimazioni di $f(x)$ soltanto in punti vicini ad x_0 . Si ricordi per esempio la serie di Mac Laurin della funzione $\arctg(x)$:

$$\arctg(x) = x - \frac{x^3}{3} + \frac{x^5}{5} + \dots + (-1)^{k+1} \frac{x^{2k-1}}{2k-1} + \dots$$

Per approssimare $\pi/4 = \arctg(1)$ con errore inferiore a 10^{-3} bisogna sommare 500 termini della serie numerica:

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots + (-1)^{k+1} \frac{1}{2k-1} + \dots$$

Ciò è comprensibile perché il polinomio di Taylor è costruito soltanto sui valori della funzione e delle sue derivate in x_0 , ed è facile vedere che in quel punto il polinomio $p_n(x)$ e le sue derivate fino all'ordine $n-1$ coincidono con la funzione $f(x)$ e le sue corrispondenti derivate. L'effetto di questo sbilanciamento è che per avere approssimazioni uniformi su intervalli sufficientemente ampi bisogna prendere polinomi di grado troppo alto.

Un modo per superare questi inconvenienti consiste nel considerare un generico polinomio

$$p_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

e determinare i suoi coefficienti attraverso altre condizioni; per esempio imponendo dei vincoli di **interpolazione** su più punti, detti **nodi d'interpolazione**, distribuiti in tutto l'intervallo sul quale si intende approssimare la funzione. Su questi nodi si impone soltanto che il polinomio coincida con la funzione, così si ha anche il vantaggio di allargare la classe di funzioni approssimabili per le quali non è più richiesta la conoscenza delle derivate ma soltanto dei valori nei nodi. In questo modo si possono anche approssimare funzioni discontinue e dati sperimentali.

L'interesse per i **polinomi algebrici** (indicheremo con π_n l'insieme dei polinomi di grado n) come possibili approssimanti di funzioni continue su intervalli chiusi e limitati è dettato dal seguente teorema:

Teorema di Weierstrass. *Assegnata una funzione $f(x)$ continua su un intervallo chiuso e limitato $[a,b]$, per ogni $\varepsilon > 0$ esiste un polinomio $p(x)$ di grado opportuno tale che $\max_{x \in [a,b]} |p(x) - f(x)| < \varepsilon$.*

Esso ci assicura che le funzioni continue su intervalli chiusi e limitati sono approssimabili uniformemente, al di sotto di qualunque tolleranza prefissata, con polinomi algebrici. Ciò si esprime anche dicendo che:

I polinomi algebrici sono densi nell'insieme $C^0[a,b]$ rispetto alla norma:

$$\|f\|_\infty := \max_{x \in [a,b]} |f(x)|$$

detta norma lagrangiana.

Dunque esiste una successione di polinomi $p_n(x)$, ciascuno di grado n , che converge uniformemente alla funzione $f(x)$. In particolare il seguente teorema fornisce, in concreto, una tale successione.

Teorema di Bernstein : *Per ogni funzione $f(t) \in C^0[0,1]$ la successione dei polinomi di Bernstein:*

$$B_n(f,t) = \sum_{i=0}^n \binom{n}{i} t^{n-i} (1-t)^i f\left(\frac{i}{n}\right)$$

converge uniformemente ad $f(t)$.

Un semplice cambio di variabile $t = \frac{x-a}{b-a}$ trasforma $[a,b]$ in $[0,1]$ e quindi consente di passare da $f(x) \in C^0[a,b]$ a $f(t) \in C^0[0,1]$. I polinomi di Bernstein per $f(x)$ sono allora:

$$B_n\left(f, \frac{x-a}{b-a}\right) = \sum_{i=0}^n \binom{n}{i} \left(\frac{x-a}{b-a}\right)^{n-i} \left(1 - \frac{x-a}{b-a}\right)^i f\left(a + \frac{i}{n}(b-a)\right)$$

I polinomi di Bernstein sono utili per tracciare curve che approssimano dati sperimentali (curve di Beziér), ma non sono sempre una buona scelta per approssimare le funzioni. In generale, a parità di grado ci sono dei polinomi che approssimano meglio la funzione data.

A questo proposito si dimostra che, per ogni $f(x) \in C^0[a,b]$, esiste uno ed un solo polinomio $p_n^*(x) \in \pi_n$ di minima distanza da $f(x)$, cioè tale che:

$$\|p_n^* - f\|_\infty \leq \|p_n - f\|_\infty \quad \forall p_n \in \pi_n$$

Il polinomio $p_n^*(x)$ è detto **polinomio di minimo scarto** per la funzione $f(x)$, e lo scarto, detto **errore minimax**, è indicato con:

$$E_n(f) := \|p_n^* - f\| = \min_{p_n \in \pi_n} (\max_{x \in [a,b]} |p_n(x) - f(x)|).$$

Per il teorema di Weierstrass si ha, ovviamente, che per ogni funzione $f(x) \in C^0[a,b]$, $E_n(f) \rightarrow 0$ per $n \rightarrow \infty$. La rapidità con cui $E_n(f)$ tende a zero, cioè il suo ordine d'infinitesimo, dipende dalla regolarità della funzione $f(x)$. La conoscenza di questo ordine d'infinitesimo è fondamentale per poter sviluppare una buona teoria dell'approssimazione. Concludiamo il paragrafo con alcune nozioni e risultati utili per stimare i termini $E_n(f)$.

Si definisce **modulo di continuità** della funzione $f(x)$ sull'intervallo $[a,b]$ la funzione

$$\omega(f, \delta) = \omega(\delta) := \sup_{|x-y| \leq \delta} |f(x) - f(y)|$$

Osserviamo che essa è una funzione non decrescente rispetto a δ , ed inoltre, per ogni $f(x) \in C^0[a,b]$ si ha: $\lim_{\delta \rightarrow 0} \omega(\delta) = 0$.

Teorema di Jackson: Per ogni k e per ogni $f(x) \in C^k[a,b]$, esistono delle costanti γ_k per le quali:

$$E_n(f) \leq \begin{cases} 6\omega\left(f, \frac{b-a}{2n}\right) & \text{per } f(x) \in C^0[a,b] \\ \gamma_k \left(\frac{b-a}{n}\right)^k & \text{per } f(x) \in C^k[a,b] \end{cases} \quad \forall n \geq k$$

Si osservi che, mentre per le funzioni di classe $C^k[a,b]$ con $k \geq 1$, $E_n(f)$ è infinitesimo di ordine k , per le funzioni continue esso è semplicemente un infinitesimo come $\omega\left(\frac{1}{n}\right)$.

Consideriamo una classe di funzioni intermedia tra $C^0[a,b]$ e $C^1[a,b]$.

Una funzione $f(x)$ si dice **holderiana** in $[a,b]$ se esiste un $M > 0$ ed un $0 < \alpha \leq 1$ tale che per ogni coppia di punti $x, y \in [a,b]$ si ha:

$$|f(x) - f(y)| < M |x - y|^\alpha.$$

Si noti che la holderianità implica la continuità uniforme e che per $\alpha = 1$ si ritrova la ben nota lipschitzianità. E' immediato verificare il seguente teorema:

Teorema 5.1 Per le funzioni holderiane si ha $\omega(\delta) < M\delta^\alpha$, in particolare per le funzioni lipschitziane $\omega(\delta)$ è infinitesimo di ordine 1.

Possiamo quindi concludere con il seguente corollario che raffina l'enunciato del teorema di Jackson.

Corollario 5.2: *Per le funzioni hölderiane, di coefficiente α , il minimo massimo errore $E_n(f)$ è un infinitesimo di ordine reale α . In particolare per le funzioni lipschitziane è un infinitesimo di ordine 1.*

Approssimazione con polinomi generalizzati.

Certe caratteristiche della funzione da approssimare possono suggerire l'opportunità di considerare altre classi di funzioni approssimanti diverse dai polinomi algebrici. Per esempio se la funzione è periodica di periodo 2π è naturale approssimarla con **polinomi trigonometrici** del tipo:

$$t_{2n+1}(x) = a_0 + a_1 \sin(x) + b_1 \cos(x) + \dots + a_n \sin(nx) + b_n \cos(nx)$$

Se invece sappiamo che $f(x)$ si annulla per $x \rightarrow \infty$ allora i seguenti **polinomi esponenziali**:

$$e_n(x) = a_0 + a_1 e^{-x} + a_2 e^{-2x} + \dots + a_n e^{-nx}$$

sono più adatti ad una sua approssimazione anche su un intervallo illimitato. Se poi la funzione $f(x)$ o una sua derivata presenta qualche discontinuità in un punto, allora i **polinomi a tratti** possono fornire dei risultati migliori. Per esempio la funzione $f(x)=|\sin(x)|$ in $[-1,1]$ è approssimabile meglio con due tratti rettilinei a sinistra ed a destra di $x=0$ che non con una parabola su tutto l'intervallo. Altre classi di funzioni approssimanti che considereremo sono le **funzioni splines**, che definiremo più avanti, e i **polinomi razionali** del tipo:

$$p_{n,m}(x) = \frac{a_0 + a_1 x + \dots + a_n x^n}{b_0 + b_1 x + \dots + b_m x^m}$$

Un teorema molto "potente" per verificare la densità di alcune classi di funzioni approssimanti in $C^0[a,b]$ è la seguente generalizzazione del teorema di Weierstrass:

Teorema di Stone-Weierstrass (1960): *Sia $A[X]$ un'algebra di funzioni definite su un insieme compatto X , cioè tale che se f e $g \in A[X]$ anche $fg \in A[X]$ e $\alpha f + \beta g \in A[X]$ per ogni coppia di numeri reali α e β . Supponiamo che l'algebra $A[X]$:*

- contenga le costanti

- separi i punti di X (cioè $\forall x, y \in X, \exists f \in A[X]$ tale che $f(x) \neq f(y)$).

Allora l'algebra $A[X]$ è densa in $C^0[X]$ nella norma lagrangiana. In altre parole per ogni funzione $f(x)$, continua su X , esiste una successione di elementi dell'algebra che converge uniformemente ad $f(x)$.

Si osservi che i polinomi algebrici costituiscono un'algebra generata dalla costante 1 e dalla funzione identica x , così come i polinomi esponenziali sono generati da 1 e da e^{-x} . Entrambe contengono le costanti e separano i punti. Più in generale, ogni insieme generato dalla funzione costante uguale ad 1 e da una funzione monotona $\varphi(x)$, forma un'algebra che contiene le costanti, separa i punti e risulta quindi densa in $C^0[a,b]$. Col precedente teorema si dimostra anche immediatamente che i polinomi di n variabili sono densi nell'insieme delle funzioni continue definite su un compatto di R^n . Per quanto riguarda i polinomi trigonometrici, vale il seguente teorema:

Il Teorema di Weierstrass. I polinomi trigonometrici

$$t_n(x) = a_0 + a_1 \sin(x) + b_1 \cos(x) + \dots + a_n \sin(nx) + b_n \cos(nx)$$

sono densi nell'insieme $C_{2\pi}$ delle funzioni continue e periodiche di periodo 2π .

Dim. E' sufficiente dimostrare la tesi sull'intervallo di periodicità $[0, 2\pi]$. La dimostrazione non può discendere direttamente dal precedente teorema di Stone-Weierstrass in quanto, pur essendo i polinomi trigonometrici una algebra, essa non separa la coppia di punti $x=0$ e $y=2\pi$. D'altra parte escludendo un estremo, l'intervallo non è più compatto. L'ostacolo si può aggirare considerando la seguente rappresentazione parametrica del circolo unitario di R^2

$$u = \sin(x); \quad v = \cos(x)$$

ed osservando che ogni polinomio trigonometrico $t_n(x)$ in $[0, 2\pi]$ può essere rappresentato con un polinomio algebrico di grado n nelle due variabili $\sin(x)$ e $\cos(x)$ e quindi nelle variabili u e v ristrette al circolo. Tale algebra separa i punti del circolo e, essendo quest'ultimo compatto in R^2 , tutte le ipotesi del teorema di Stone-Weierstrass sono verificate.

Una volta scelto il tipo di funzione con la quale approssimare la funzione data, si tratta di stabilire dei criteri per la sua determinazione. Questi criteri sono dati sostanzialmente da due tipi di vincoli cui devono soddisfare i parametri che definiscono la

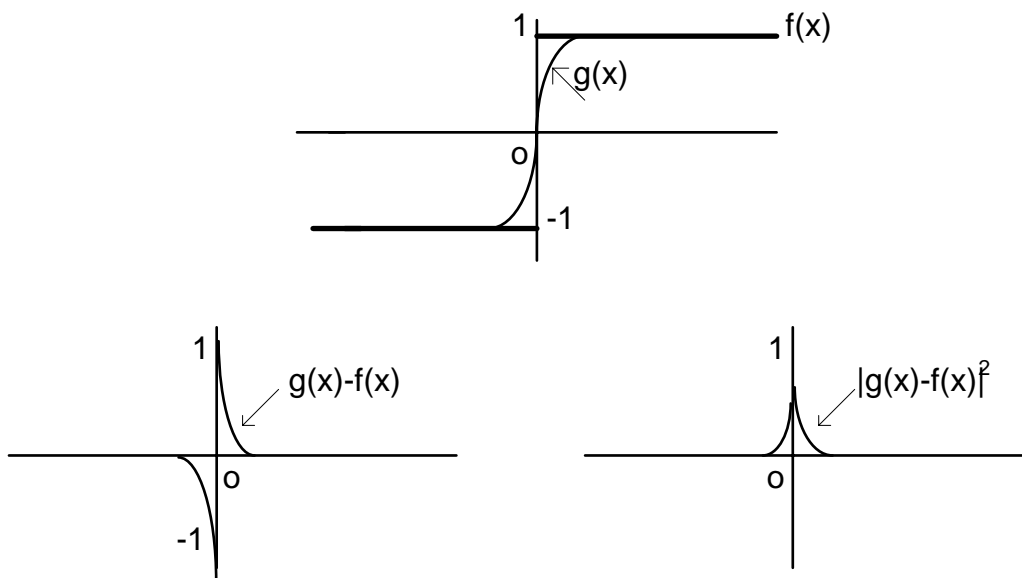
funzione approssimante. Abbiamo già accennato ai **vincoli interpolatori** con i quali si impone che in alcuni punti i valori della funzione approssimante e, più in generale, di certe sue derivate, coincidano con quelli della funzione data. Un secondo tipo di condizioni sono date dai **vincoli variazionali** con i quali si impone che la funzione approssimante sia la più prossima alla funzione data in qualche norma atta a valutare la distanza tra due funzioni. Tra le varie norme possibili, consideriamo la **norma lagrangiana**:

$$\|f - g\|_{\infty} = \sup_{x \in [a,b]} |f(x) - g(x)|$$

e la **norma hilbertiana**:

$$\|f - g\|_2 = \sqrt{\int_a^b (f(x) - g(x))^2 dx}.$$

La norma hilbertiana è particolarmente adatta a misurare la distanza tra funzioni discontinue. Nelle seguenti figure si vede che la distanza tra la funzione $g(x)$ e la funzione (discontinua) $f(x)$ è 1 in norma lagrangiana mentre è molto piccola in norma hilbertiana



Si osservi che, nell'esempio proposto, nessuna funzione continua può avere distanza lagrangiana da $f(x)$ minore di 1. In particolare la funzione nulla approssima la $f(x)$ con lo stesso errore della $g(x)$, il che non è ragionevole. Ciò rende la distanza lagrangiana inappropriata all'esempio considerato, mentre la distanza hilbertiana consente delle stime più sensate.

Una terza classe di norme, dette **norme hilbertiane discrete** sono del tipo:

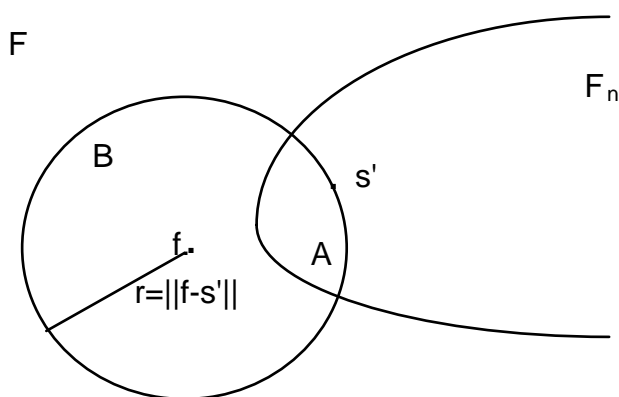
$$\|f(x)\|_{2,d} = \sqrt{\sum_{i=1}^m w_i f(x_i)^2}$$

e, come vedremo, consentono di individuare un legame tra i metodi variazionali e quelli interpolatori.

Abbiamo già accennato al fatto che per ogni funzione $f(x)$ esiste un polinomio di grado n di minima distanza lagrangiana da $f(x)$. Estendiamo questo risultato al caso dei polinomi generalizzati ed al caso di una generica norma $\| \cdot \|$.

Teorema della miglior approssimazione. *Sia F uno spazio lineare normato e F_n un sottospazio di dimensione n generato dalla base s_1, \dots, s_n . Per ogni elemento $f \in F$, esiste un elemento $s^* = a_1 s_1 + \dots + a_n s_n \in F_n$, di minima distanza da f tra tutti quelli di F_n , cioè tale che $\|f - s^*\| \leq \|f - s\| \quad \forall s \in F_n$.*

Dim. Sia s' un elemento qualunque di F_n e sia B la palla di centro f e raggio $r = \|f - s'\|$.



L'insieme $A := B \cap F_n$ è limitato e chiuso, in quanto intersezione di chiusi. Essendo inoltre incluso in un sottospazio a dimensione finita, risulta compatto. La funzione $\varphi(s) = \|f - s\|: F \rightarrow \mathbb{R}^+ \cup \{0\}$ è una funzione continua e quindi ammette un punto minimo s^* nel compatto A . Ogni altro punto di F_n ha distanza da f maggiore di $\|f - s^*\|$ e quindi s^* è l'elemento di minimo scarto da f in tutto F_n .

Ricordiamo infine che la norma hilbertiana è dedotta dal prodotto scalare:

$$\langle f | g \rangle = \int_a^b f(x)g(x)dx$$

il quale dà luogo alla nozione di ortogonalità tra funzioni. Analogamente a quanto visto nel caso degli spazi \mathbb{R}^n , ciò consentirà di risolvere in modo relativamente facile il problema della ricerca dell'elemento di minima distanza.

2. METODI INTERPOLATORI.

Interpolazione di Lagrange.

Data una funzione $f(x)$ definita in un intervallo $[a,b]$ ed un insieme di nodi $\Delta_n = \{x_0 < x_1 < \dots < x_n\}$ in $[a,b]$, si definisce **polinomio d'interpolazione di Lagrange** il polinomio $p_n(x) \in \pi_n$ che soddisfa le condizioni d'interpolazione:

$$p_n(x_i) = f(x_i) \quad i=0,1,\dots,n.$$

Tale polinomio esiste ed è unico per ogni insieme Δ_n di nodi (tra loro distinti) e per ogni funzione $f(x)$. Infatti, esprimendo $p_n(x)$ in *forma canonica*:

$$p_n(x) = a_0 + a_1x + \dots + a_nx^n,$$

le condizioni di interpolazione si traducono nel seguente sistema lineare nelle incognite a_0, a_1, \dots, a_n :

$$a_0 + a_1x_0 + \dots + a_nx_0^n = f(x_0)$$

$$a_0 + a_1x_1 + \dots + a_nx_1^n = f(x_1)$$

.....

$$a_0 + a_1x_n + \dots + a_nx_n^n = f(x_n).$$

Il determinante della matrice del sistema, detto *determinante di Vandermonde*, è:

$$V(x_0, x_1, \dots, x_n) = \begin{vmatrix} 1 & x_0 & \dots & x_0^n \\ 1 & x_1 & \dots & x_1^n \\ \dots & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^n \end{vmatrix}$$

ed è non nullo per ogni insieme di nodi distinti. Il sistema ammette quindi una ed una sola soluzione.

A volte sarà conveniente indicare con $L_{\Delta_n}(f,x)$ l'operatore che associa alla funzione $f(x)$ il suo polinomio d'interpolazione sui nodi Δ_n . Per l'unicità dell'interpolazione (o, se si

preferisce, per il *principio di identità dei polinomi*), è chiaro che per ogni Δ_n e per ogni polinomio $q_n \in \pi_n$ si ha:

$$L_{\Delta_n}(q_n, x) = q_n(x)$$

Sul piano pratico non conviene rappresentare il polinomio d'interpolazione nella forma canonica, ma attraverso i seguenti *coefficienti di Lagrange* $\ell_i(x)$:

$$\ell_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j} = \frac{(x - x_0)(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)} \in \pi_n$$

per i quali si ha:

$$\ell_i(x_j) = \delta_{i,j} = \begin{cases} 0 & \text{per } i \neq j \\ 1 & \text{per } i = j \end{cases}$$

Si osservi che essi dipendono esclusivamente dai nodi e sono tutti polinomi di grado n .

Il polinomio d'interpolazione assume quindi la forma:

$$L_{\Delta_n}(f, x) = \sum_{i=0}^n \ell_i(x) f(x_i)$$

dalla quale appare evidente che l'operatore $L_{\Delta_n}(f, x)$ è lineare rispetto ad f . Il polinomio d'interpolazione scritto in questa forma lo chiameremo **polinomio di Lagrange**.

Errore dell'interpolazione di Lagrange.

Per dare una valutazione puntuale dell'errore di interpolazione $f(x) - p_n(x)$ consideriamo un punto $\bar{x} \in [a, b]$, diverso dai nodi Δ_n , e con esso costruiamo la funzione:

$$g(x) = f(x) - p_n(x) - \frac{w(x)}{w(\bar{x})} (f(\bar{x}) - p_n(\bar{x}))$$

dove

$$w(x) = (x - x_0)(x - x_1) \dots (x - x_n).$$

Supponiamo che la funzione $f(x)$, e quindi la $g(x)$, sia di classe $C^{n+1}[a, b]$. Poiché $g(x)$ si annulla su $n+2$ punti, i nodi ed il punto \bar{x} , la sua derivata $g'(x)$ si annulla in almeno $n+1$ punti interni all'intervallo $[a, b]$. Così la derivata seconda $g''(x)$ si annulla in almeno n

punti e, così proseguendo, la derivata $g^{(n+1)}(x)$ si annulla in almeno un punto interno di $[a,b]$. Detto $\bar{\xi}$ tale punto, che risulta dipendere da \bar{x} , per esso si ha:

$$g^{(n+1)}(\bar{\xi}) = f^{(n+1)}(\bar{\xi}) - \frac{(n+1)!}{w(\bar{x})} (f(\bar{x}) - p_n(\bar{x})) = 0.$$

e quindi $\forall \bar{x} \in \Delta_n$

$$f(\bar{x}) - p_n(\bar{x}) = \frac{f^{(n+1)}(\bar{\xi}) w(\bar{x})}{(n+1)!}.$$

Poichè sui nodi il polinomio $w(x)$ si annulla, la precedente stima puntuale è valida per ogni punto di $[a,b]$. Non conoscendo il valore $\bar{\xi}$ in funzione di \bar{x} , la formula può essere utilizzata solo maggiorando $f^{(n+1)}(x)$ su tutto $[a,b]$:

$$|f(\bar{x}) - p_n(\bar{x})| \leq \frac{\|f^{(n+1)}\|_{\infty} |w(\bar{x})|}{(n+1)!}$$

Da quest'ultima si può infine ricavare la stima uniforme:

$$\|f - p_n\|_{\infty} \leq \frac{\|f^{(n+1)}\|_{\infty} \|w\|_{\infty}}{(n+1)!}. \quad (5.1)$$

Si osservi che la stima dell'errore dipende da due fattori indipendenti tra loro. Il primo, $\|f^{(n+1)}\|_{\infty}$, dipende soltanto dalla regolarità della funzione $f(x)$; il secondo $\|w\|_{\infty}$, dipende solo dai nodi. E' facile vedere che una maggiorazione brutale del termine $\|w\|_{\infty}$ è data da $(n+1)!h^{n+1}$, dove h è la massima distanza tra due punti consecutivi dell'insieme $\{a, x_0, x_1, \dots, x_n, b\}$. Vale allora la stima:

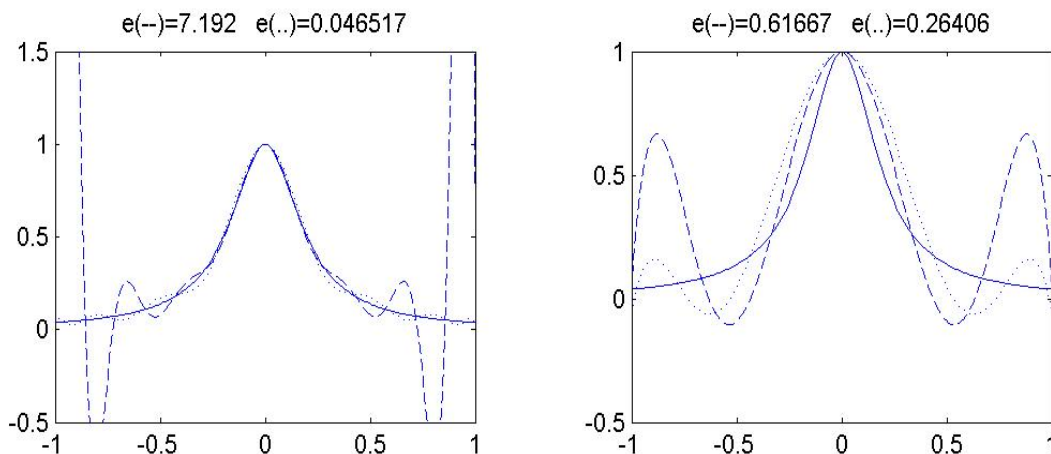
$$\|f - p_n\|_{\infty} \leq \|f^{(n+1)}\|_{\infty} h^{n+1} \quad (5.2)$$

Se, in particolare, $x_0=a$ e $x_n=b$, cioè se ci limitiamo a dare una stima dell'errore all'interno dei nodi di interpolazione, allora il termine $\|w\|_{[x_0, x_n]} = \max_{[x_0, x_n]} |w(x)|$ può essere maggiorato con $\frac{1}{4}(n!)h^{n+1}$, dove $h = \max_{1 \leq i \leq n} (x_i - x_{i-1})$. In questo caso vale la seguente maggiorazione:

$$\|f - p_n\|_{[x_0, x_n]} \leq \frac{\|f^{(n+1)}\|_{[x_0, x_n]} h^{n+1}}{4(n+1)} \quad (5.3)$$

Se i nodi sono equidistanti, le precedenti stime di $\|w\|_\infty$ sono le migliori possibili e quindi (5.2) e (5.3) sono ottimali per aumentare l'errore di interpolazione. A differenza di quello che si potrebbe pensare, la distribuzione di nodi equidistanti non è però la migliore nell'interpolazione con polinomi algebrici. Vedremo infatti che il termine $\|w\|_\infty$ è minimizzato se i nodi sono gli zeri dei polinomi di Chebyshev relativi all'intervallo $[a,b]$. E' tipico della interpolazione su nodi equidistanti il seguente *fenomeno di Runge* dove, al crescere di n , i polinomi interpolanti aumentano l'oscillazione.

Nelle seguenti figure si interpola la funzione $f(x)=\frac{1}{1+x^2}$ su 7 e 15 nodi equidistanti (linea tratteggiata) e di Chebyshev (linea punteggiata).



Polinomi di Chebyshev.

Per ogni numero naturale n si definisce **polinomio di Chebyshev di grado n** la seguente funzione definita in $[-1,1]$:

$$T_n(x) = \cos(n \arccos(x)).$$

Osserviamo innanzitutto che $\forall n, T_n(x) \in \pi_n$. Per le note formule sulla somma del coseno si ha:

$$\cos(a) + \cos(b) = 2 \cos\left(\frac{a+b}{2}\right) \cos\left(\frac{a-b}{2}\right).$$

le quali, attraverso il cambio di variabile

$$\varphi = \arccos(x) \quad -1 \leq x \leq 1,$$

forniscono:

$$\cos((n+1)\varphi) + \cos((n-1)\varphi) = 2 \cos(n\varphi) \cos(\varphi)$$

e quindi:

$$T_{n+1}(x) + T_{n-1}(x) = 2T_n(x) x.$$

Si ottiene così la **relazione ricorsiva** a tre termini:

$$T_{n+1}(x) = 2x T_n(x) - T_{n-1}(x) \quad n=1,2,\dots$$

Poichè $T_0(x)=1$ e $T_1(x)=x$, si vede facilmente che $T_n(x) \in \pi_n$ per ogni n , ed inoltre il **coefficiente principale** di $T_n(x)$ è $k_n=2^{n-1}$.

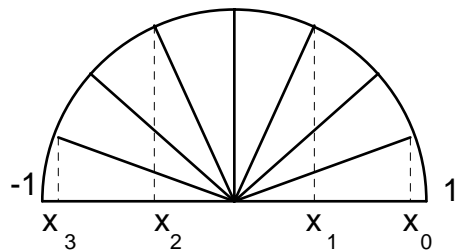
Il polinomio $T_n(x)$ è ovviamente definito su tutto \mathbb{R} , ma soltanto in $[-1,1]$ coincide con la funzione $\cos(n \arccos(x))$.

Verifichiamo ora che, per ogni n , $T_n(x)$ ammette n radici reali e distinte in $(-1,1)$. Poichè il coseno si annulla sui punti $z_k = (2k-1)\frac{\pi}{2}$ con $k \in \mathbb{Z}$, la funzione $\cos(n \arccos(x))$ si annullerà sui punti x_k tali che $n \arccos(x_k) = z_k$. Quindi, restringendoci ai valori dell'arccoseno in $[0, \pi]$,

$$\arccos(x_k) = \frac{(2k-1)\pi}{2n} \quad k=1,2,\dots,n$$

e

$$x_k = \cos\left(\frac{(2k-1)\pi}{2n}\right) \quad k=1,2,\dots,n.$$



radici di $T_4(x)$

Per quanto riguarda i **massimi e minimi** di $T_n(x)$, essi sono raggiunti in quei punti y_k per i quali $\arccos(y_k) = k\pi/n$ con $k \in \mathbb{Z}$. Restringendoci ancora ai valori dell'arccoseno in $[0, \pi]$, si ha:

$$\arccos(y_k) = \frac{k\pi}{n} \quad k=0,1,\dots,n$$

e quindi:

$$y_k = \cos \frac{k\pi}{n} \quad k=0,1,\dots,n.$$

Su tali punti estremali si ha:

$$T_n(y_k) = (-1)^k \quad k=0,1,\dots,n.$$

Sia ora $\tilde{T}_n(x) = \frac{1}{2^{n-1}} T_n(x)$ il **polinomio monico di Chebyshev** la cui norma è:

$$\|\tilde{T}_n(x)\| = \frac{1}{2^{n-1}} \|T_n(x)\| = \frac{1}{2^{n-1}} \quad (5.4)$$

e dimostriamo per esso la seguente proprietà di minimo.

Teorema 5.3: *Tra tutti i polinomi monici di grado n , il polinomio $\tilde{T}_n(x)$ possiede la minima norma lagrangiana nell'intervallo $[-1, 1]$.*

Dim: Supponiamo che esista un polinomio monico $p_n(x)$ di norma inferiore a $\tilde{T}_n(x)$

$$\|p_n(x)\| < \|\tilde{T}_n(x)\|.$$

Poichè i massimi e minimi di $\tilde{T}_n(x)$ si alternano sui punti estremali, si avrà:

$$\tilde{T}_n(y_0) > p_n(y_0)$$

$$\tilde{T}_n(y_1) < p_n(y_1)$$

....

e così di seguito, in modo alternato, fino al punto y_n .

Il polinomio $\tilde{T}_n(x) - p_n(x)$ cambia segno $n+1$ volte e quindi possiede n zeri. Essendo differenza di due polinomi monici di grado n , esso è di grado al più $n-1$ e quindi coincide con la funzione nulla. I due polinomi sono dunque uguali ma questo

contraddice la supposizione che abbiano norma diversa. In conclusione non esiste polinomio monico di norma inferiore alla norma di $\tilde{T}_n(x)$.

Questo risultato può essere utilizzato, in riferimento alla formula (5.1), per la scelta ottimale dei nodi di interpolazione in un generico intervallo $[a,b]$. Consideriamo a tale scopo il cambio di variabile

$$x = \frac{1}{2}(b+a) + \frac{1}{2}(b-a)t \quad t \in [-1,1]$$

che trasforma $[-1,1]$ in $[a,b]$. Il polinomio $w(x)=(x-x_0)(x-x_1)\dots(x-x_n)$ diventa:

$$w(x)=w'(t)=\frac{1}{2^{n+1}}(b-a)^{n+1}(t-t_0)(t-t_1)\dots(t-t_n)$$

e quindi:

$$\|w(x)\|_{[a,b]} = \|w'(t)\|_{[-1,1]} = \frac{1}{2^{n+1}}(b-a)^{n+1} \|(t-t_0)(t-t_1)\dots(t-t_n)\|_{[-1,1]} \quad (5.5)$$

Poichè il termine $\|(t-t_0)(t-t_1)\dots(t-t_n)\|_{[-1,1]}$ risulta minimo quando i punti t_k sono gli zeri di $\mathcal{T}'_{n+1}(t)$, il termine $\|w(x)\|_{[a,b]}$ sarà minimo sui punti

$$x_k = \frac{1}{2}(b+a) + \frac{1}{2}(b-a)t_k \quad t_{k-1} = \cos\left(\frac{(2k-1)\pi}{n+1}\right) \quad k=1,2,\dots,n+1.$$

Per la (5.4) e (5.5) si avrà infine:

$$\|w(x)\|_{[a,b]} = \frac{1}{2^{n+1}}(b-a)^{n+1} \frac{1}{2^n} = \frac{1}{2 \cdot 4^n}(b-a)^{n+1}$$

Convergenza dell'interpolazione di Lagrange:

Consideriamo la successione $\{p_n(x)\}_{n=0}^{\infty}$ di polinomi d'interpolazione ottenuta in corrispondenza ad una successione di insiemi di nodi $\{\Delta_n\}_{n=0}^{\infty}$, tutti inclusi in $[a,b]$, per i quali supporremo semplicemente che la massima distanza di due nodi successivi di Δ_n , che ora chiameremo h_n , tenda a zero al crescere di n .

Nel caso che f sia di classe $C^\infty[a,b]$, le formule (5.1), (5.2) e (5.3) forniscono implicitamente le condizioni per la convergenza uniforme della successione di polinomi

interpolanti alla funzione $f(x)$. Se, per esempio, le derivate di f sono equilimitate, cioè se $\|f^{n+1}\|_\infty < M \quad \forall n$, allora è evidente che $p_n(x)$ converge uniformemente ad $f(x)$, ma ciò può accadere in condizioni molto più generali. E' da osservare però che, come abbiamo visto in relazione al fenomeno di Runge, sui nodi equidistanti la convergenza non è garantita per tutte le funzioni di classe C^∞ .

D'altra parte vedremo che sugli zeri dei polinomi di Chebyshev la convergenza dei polinomi d'interpolazione è garantita per tutte le funzioni holderiane su $[a,b]$ ma la sola continuità non è sufficiente per nessuna scelta dei nodi. Vale infatti il seguente teorema.

Teorema di Faber. *Per ogni successione $\{\Delta_n\}_{n=0}^\infty$ di nodi in $[a,b]$, esiste una funzione $f(x) \in C^\circ[a,b]$ per la quale la successione di polinomi d'interpolazione non converge.*

Per dare una stima dell'errore di interpolazione nel caso generale di funzioni che siano soltanto continue, dimostriamo il seguente teorema:

Teorema 5.4: *Per ogni insieme di nodi Δ_n di $[a,b]$ e per ogni $f \in C^\circ[a,b]$ si ha la seguente stima dell'errore d'interpolazione*

$$\|f - L_{\Delta_n}(f)\|_\infty \leq E_n(f) (1 + \lambda_n)$$

dove il termine:

$$\lambda_n = \max_{x \in [a,b]} \sum_{i=0}^n |\ell_i(x)|$$

è detto **numero di Lebesgue** associato a Δ_n .

Dim. Sia $p_n^*(x)$ il polinomio di minimo scarto per $f(x)$ in $[a,b]$. Poichè $p_n^* = L_{\Delta_n}(p_n^*)$ si ha:

$$\begin{aligned} \|f - L_{\Delta_n}(f)\|_\infty &= \|f - p_n^* + L_{\Delta_n}(p_n^*) - L_{\Delta_n}(f)\|_\infty \leq \|f - p_n^*\| + \|L_{\Delta_n}(p_n^*) - L_{\Delta_n}(f)\|_\infty \\ &\leq E_n(f) + \|L_{\Delta_n}(p_n^* - f)\|_\infty \leq E_n(f) + \left\| \sum_{i=0}^n \ell_i(x) (p_n^*(x_i) - f(x_i)) \right\|_\infty \\ &\leq E_n(f) + \max_{x \in [a,b]} \left| \sum_{i=0}^n \ell_i(x) (p_n^*(x_i) - f(x_i)) \right| \\ &\leq E_n(f) + \max_{x \in [a,b]} \sum_{i=0}^n |\ell_i(x)| |p_n^*(x_i) - f(x_i)| \\ &\leq E_n(f) + E_n(f) \left(\max_{x \in [a,b]} \sum_{i=0}^n |\ell_i(x)| \right) = E_n(f) (1 + \lambda_n). \end{aligned}$$

Sappiamo che il minimo massimo errore $E_n(f)$ è infinitesimo per ogni funzione continua, quindi la possibilità che una successione d'interpolanti converga dipende dal prodotto $E_n(f) \lambda_n$. Sfortunatamente non c'è nessuna successione di nodi $\{\Delta_n\}_{n=0}^{\infty}$ per la quale la corrispondente successione $\{\lambda_n\}_{n=0}^{\infty}$ di numeri di Lebesgue sia limitata. Se ciò fosse vero la convergenza su quei nodi sarebbe assicurata per ogni funzione continua e ciò sarebbe in contraddizione con il teorema di Faber. Sarebbe interessante sapere, per ogni successione di nodi, con quale ordine divergono i corrispondenti numeri λ_n così da contrapporre una successione $E_n(f)$ per la quale il prodotto $E_n(f) \lambda_n$ rimanga infinitesimo. A questo proposito vale il seguente teorema:

Teorema di Natanson. *Per ogni successione di nodi, i corrispondenti numeri di Lebesgue soddisfano la relazione:*

$$\lambda_n > \frac{2}{\pi} \log(n) - c$$

In particolare sui nodi di Chebyshev si ha:

$$\lambda_n < \frac{2}{\pi} \log(n) + c.$$

Da questo teorema ricaviamo ancora una volta il risultato che i nodi di Chebyshev sono *ottimali* per l'interpolazione. Concludiamo il paragrafo con il seguente teorema che ci dice qual'è la regolarità richiesta alla funzione f affinché, sui nodi di Chebyshev, il prodotto $E_n(f) \lambda_n$ sia convergente.

Teorema 5.5: *Per ogni funzione hölderiana su $[a,b]$ l'interpolazione sui nodi di Chebyshev è convergente.*

Dim: Sappiamo che per ogni funzione continua sia ha la stima:

$$E_n(f) \leq 6\omega\left(f, \frac{b-a}{2n}\right) = 6 \sup_{|x-y| \leq \frac{b-a}{2n}} |f(x) - f(y)|.$$

D'altra parte f è hölderiana e quindi, per il teorema 5.1,

$$E_n(f) \leq 6\omega\left(f, \frac{b-a}{2n}\right) \leq 6M\left(\frac{b-a}{2n}\right)^\alpha$$

In conclusione si ha:

$$E_n(f)_{\lambda_n} \leq 6M \left(\frac{b-a}{2n} \right)^\alpha \left(\frac{2}{\pi} \log(n) + c \right)$$

che risulta infinitesimo per ogni $\alpha > 0$. La convergenza è quindi dimostrata in base al teorema 5.4.