

Analisi Numerica

AlpT (@freaknet.org)

<http://freaknet.org/alpt>

April 21, 2010

Abstract

Questo testo e' una rielaborazione personale degli appunti presi durante il corso di Analisi Numerica, tenuto dal Prof. Giovanni Russo¹ presso il dipartimento di Matematica, Catania, A.A. 2007/2008.

Saro' ben lieto di correggere ogni eventuale errore che mi comunicherai.

Buon lettura.

^ ^

¹Sito web del Prof. Russo: <http://www.dmi.unict.it/~russo/>

Copyright ©2008 Andrea Lo Pumo aka AlpT <alpt@freaknet.org>. All rights reserved.

This document is free; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

This document is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this document; if not, write to the Free Software Foundation, Inc., 675 Mass Ave, Cambridge, MA 02139, USA.

Contents

1	ChangeLog / Correzioni / Errata	1
2	Proposizioni varie	2
3	Preliminari	3
4	Norme	4
5	Sistemi lineari	13
5.1	Fattorizzazione LU	19
5.1.1	Applicazioni	25
5.2	Fattorizzazione di Cholesky	27
5.3	Autovalori	28
6	Interpolazione e approssimazione	34
6.1	Interpolazione polinomiale	35
7	Approssimazione polinomiale	44
7.1	Spline	46
8	Metodo dei minimi quadrati	49
9	Equazioni non lineari	54
9.1	Ricerca di punti fissi	58
9.2	Ordine di convergenza	62
9.2.1	Criteri di arresto	64
9.3	Sistemi non lineari	65
9.4	Massimi e minimi	66
10	Integrali	68
10.1	Formule composite	77
11	Formulario	80
12	*	83

1 ChangeLog / Correzioni / Errata

Cristina Oliveri ha segnalato alcune correzioni che potete trovare qui: <http://www.freaknet.org/alpt/math/appunti/PDF/analisi/numerica/analisi-numerica-patch-Cristina-Oliveri.pdf>.

2 Proposizioni varie

Proposition 2.1.

$$\sum_{i=1}^n i = \frac{n(n+1)}{2}$$

Proof: Supponiamo dapprima n dispari:

$$\begin{aligned} \sum_{i=1}^n i &= 1 + 2 + 3 + \cdots + (n-3) + (n-2) + (n-1) + n = \left(1 + (2 + (3 + \cdots + (n-3) + n-2) + n-1)\right) + n = \\ &= \underbrace{(n + n + n + \cdots + n)}_{(n-1)/2 \text{ volte}} + n = \frac{n-1}{2}n + n = \frac{n(n+1)}{2} \end{aligned}$$

Se n e' pari:

$$\sum_{i=1}^n i = n + \sum_{i=1}^{n-1} i \underbrace{=}_{n-1 \text{ dispari}} n + \frac{(n-1)n}{2} = \frac{2n + n^2 - n}{2} \frac{n(n+1)}{2}$$

□

Proposition 2.2. *Somma quadrati:*

$$\sum_{i=1}^n i^2 = \frac{2n^3 + 3n^2 + n}{6}$$

Proof: Data una successione $\{a_i\}$, poniamo:

$$\Delta a_i = a_{i+1} - a_i$$

Si ha

$$\Delta i^3 = (i+1)^3 - i^3 = 1 + 3i^2 + 3i$$

$$3i^2 = \Delta i^3 - 3i - 1$$

$$3 \sum_{i=1}^n i^2 = \sum_{i=1}^n (\Delta i^3 - 3i - 1) =$$

$$\sum_{i=1}^n \Delta i^3 = \sum_{i=1}^n (i+1)^3 - i^3 = 2^3 - 1^3 + 3^3 - 2^3 + \cdots + (n+1)^3 - n^3 = (n+1)^3 - 1^3$$

$$= (n+1)^3 - 1^3 - 3 \sum_{i=1}^n i - \sum_{i=1}^n 1 \underbrace{=}_{[2.1, \text{pg.2}]} (n+1)^3 - 1^3 - 3 \frac{n(n+1)}{2} - n$$

$$\sum_{i=1}^n i^2 = \frac{2(n+1)^3 - 2 - 3n(n+1) - 2n}{6} = \frac{2n^3 + 2 + 6n^2 + 6n - 2 - 3n^2 - 3n - 2n}{6} = \frac{2n^3 + 3n^2 + n}{6}$$

□

3 Preliminari

Definition 3.1. Definiamo il costo di un algoritmo numerico, misurato in *flops* (Floating Point Operations), come il numero di divisioni e moltiplicazioni che l'algoritmo esegue.

Useremo la seguente scrittura per il calcolo del costo di un algoritmo:

1. Il costo della procedura α , verra' indicato con $[\alpha]$
2. Verranno evidenziate le operazioni usate con degli opportuni simboli (operazione $*$, divisione $/$)
3. Ogni operazione verra' sostituita da un 1. Tutte le non-operazioni verranno eliminate

Definizioni:

$$\begin{aligned} [a * b] &\stackrel{\text{def}}{=} [a/b] = 1 \\ [a \pm b] &\stackrel{\text{def}}{=} 0 \end{aligned}$$

Come esempio completo, vedi il calcolo del costo dell'algoritmo [5.1,pg.13].

Proposition 3.2. Siano $m, n \in \mathbb{N}$, $a, b, \lambda \in \mathbb{R}$, $X \in \mathbb{R}^p$, si ha

$$\begin{aligned} \left[\sum_{i=m}^n x_i \right] &= \sum_{i=m}^n [x_i] \\ [n!] &= n - 1 \\ [(a+b)^2] &= 1 \\ [\lambda X] &= p \end{aligned}$$

Definition 3.3. Date due successioni $\{x_n\}, \{y_n\}$, poniamo

$$x_n = O(y_n) \stackrel{\text{def}}{\Leftrightarrow} \exists C, N \in \mathbb{R} : |x_n| \leq C|y_n| \quad \forall n > N$$

Usiamo la seguente convenzione:

$$a = b + O(y_n) \stackrel{\text{def}}{\Leftrightarrow} a - b = O(y_n)$$

Example 3.4. Preso $k > 1$

$$\begin{aligned} \frac{n^3}{k} + n &= O(n^3), \text{ infatti,} \\ \frac{n^3}{k} + n &\leq n^3 + n \leq 2n^3 \quad \forall n \geq 1 \\ \frac{n^3}{k} + n &= O\left(\frac{n^3}{k}\right), \text{ infatti,} \\ \frac{n^3}{k} + n &\leq 2\frac{n^3}{k} \quad \forall n \geq \sqrt{k} \end{aligned}$$

Definition 3.5. Date due funzioni $f(x), g(x) : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}$, dato $x_0 \in I$, poniamo

$$f(x) = O_{x_0}(g(x)) \stackrel{\text{def}}{\Leftrightarrow} \exists C, r \in \mathbb{R}, r > 0 : |f(x)| \leq C|g(x)| \quad \forall x \in]x_0 - r, x_0 + r[$$

Adottiamo la seguente convenzione:

$$f(x) = O(g(x)) \stackrel{\text{def}}{\Leftrightarrow} f(x) = O_0(g(x))$$

Definition 3.6. Date due funzioni $f(x), g(x) : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}$, dato $x_0 \in I$, poniamo

$$f(x) = o_{x_0}(g(x)) \stackrel{\text{def}}{\Leftrightarrow} \exists r > 0, \exists k(x) :]x_0 - r, x_0 + r[\longrightarrow \mathbb{R}_+ \text{ tale che } \begin{cases} |f(x)| \leq k(x)|g(x)| \quad \forall x \in]x_0 - r, x_0 + r[\\ \lim_{x \rightarrow x_0} k(x) = 0 \end{cases}$$

Adottiamo la seguente convenzione:

$$f(x) = o(g(x)) \stackrel{\text{def}}{\Leftrightarrow} f(x) = o_0(g(x))$$

4 Norme

Proposition 4.1. Consideriamo lo spazio vettoriale $\mathbb{R}^n$ o $\mathbb{C}^n$ e sia $v = (v_1, \dots, v_n)$ un suo vettore. Si puo' dimostrare che le seguenti funzioni sono delle norme:

$$\|v\|_1 = \sum_{i=1}^n |v_i|$$

$$\|v\|_2 = \sqrt{\sum_{i=1}^n |v_i|^2}$$

$$\|v\|_3 = \left(\sum_{i=1}^n |v_i|^3 \right)^{\frac{1}{3}}$$

$\vdots$

$$\|v\|_p = \left(\sum_{i=1}^n |v_i|^p \right)^{\frac{1}{p}}$$

$$\|v\|_\infty = \max\{|v_i| \mid i = 1, \dots, n\}$$

Proposition 4.2.

$$\lim_{p \rightarrow +\infty} \|v\|_p = \|v\|_\infty$$

Proof: Sia

$$|v_k| = \max\{|v_i| \mid i = 1, \dots, n\} = \|v\|_\infty$$

$$\Rightarrow |v_k| \geq |v_i| \quad \forall i \Leftrightarrow 1 \geq \frac{|v_i|}{|v_k|} \geq 0 \quad \forall i$$

$$\Rightarrow 1 \geq \underbrace{\lim_{p \rightarrow +\infty} \left(\frac{|v_i|}{|v_k|} \right)^p}_{s_i} \geq 0$$

$$s := \lim \sum_{i=1}^n \left(\frac{|v_i|}{|v_k|} \right)^p = \sum_{i=1}^n s_i \quad (1)$$

Se $|v_k| = 0$, la tesi e' vera. Supponiamo che $|v_k| \neq 0$.
Si ha:

$$\|v\|_p = \left(\sum_{i=1}^n |v_i|^p \right)^{\frac{1}{p}} = \left(|v_k|^p \sum_{i=1}^n \left(\frac{|v_i|}{|v_k|} \right)^p \right)^{\frac{1}{p}} = |v_k| \left(\sum_{i=1}^n \left(\frac{|v_i|}{|v_k|} \right)^p \right)^{\frac{1}{p}} \quad (2)$$

$$\lim_{p \rightarrow +\infty} \|v\|_p = \lim_{p \rightarrow +\infty} |v_k| s^{\frac{1}{p}} = |v_k| s^0 = |v_k|$$

□

Definition 4.3. Siano $\| - \|$, $\| - \|'$ due norme. Poniamo

$$\| - \| \sim \| - \|' \Leftrightarrow \exists m, M > 0 : m\|v\| \leq \|v\|' \leq M\|v\| \quad \forall v \in \mathbf{V}$$

Proposition 4.4. La relazione definita in [4.3,pg.5] e' una relazione d'equivalenza.

Proof:

1. E' riflessiva: $N(-) \sim N(-)$, infatti,

$$N(v) \leq N(v) \leq N(v) \quad \forall v \in \mathbf{V}$$

2. E' simmetrica: $N(-) \sim N(-)' \Rightarrow N(-)' \sim N(-)$, infatti,

$$N(-) \sim N(-)' \Leftrightarrow mN(v) \leq N(v)' \leq MN(v) \Rightarrow \frac{1}{M}N(v)' \leq N(v) \leq \frac{1}{m}N(v)'$$

3. E' transitiva: $N(-) \sim N(-)' \sim N(-)'' \Rightarrow N(-) \sim N(-)''$, infatti,

$$N(-) \sim N(-)' \Leftrightarrow mN(v) \leq N(v)' \leq MN(v)$$

$$N(-)' \sim N(-)'' \Leftrightarrow bN(v)' \leq N(v)'' \leq BN(v)'$$

$$bmN(v) \leq bN(v)' \leq N(v)'' \leq BN(v)' \leq BMN(v) \Leftrightarrow N(-) \sim N(-)''$$

□

Theorem 4.5. Tutte le norme in uno spazio vettoriale $\mathbf{V}$ di dimensione finita n sono equivalenti.

Proof:

$$\dim \mathbf{V} = n \Rightarrow \exists B = \{e_1, \dots, e_n\} \text{ base di } \mathbf{V}$$

Definiamo la funzione:

$$N_0 : \mathbf{V} \longrightarrow \mathbb{R}_+$$

$$N_0(v) = \max \{|v_i| \mid i = 1, \dots, n\}$$

dove con v_i abbiamo indicato la i -esima componente di v secondo B , ovvero:

$$v_i = ([v]_B)_i$$

(1) N_0 e' una norma

$$1. N_0(v) \geq 0$$

$$2. N_0(v) = 0 \Leftrightarrow v = \underline{0}$$

$$3. N_0(kv) = |k|N_0(v), \text{ infatti,}$$

$$N_0(kv) = \max \{|[kv]_B|_i \mid i = 1, \dots, n\} = |k| \max \{|[v]_B|_i \mid i = 1, \dots, n\} = |k|N_0(v)$$

$$4. N_0(v+w) \leq N_0(v) + N_0(w), \text{ infatti,}$$

$$N_0(v+w) = \max \{|(v+w)_i| \mid i = 1, \dots, n\} = \max \{|v_i + w_i| \mid i = 1, \dots, n\} \leq |v_k| + |w_k| \leq N_0(v) + N_0(w)$$

(2) Dimostriamo che una qualsiasi norma N e' equivalente a N_0 . Cosi' avremo acquisito la tesi.

Basta dimostrare che

$$\exists m, M > 0 : mN_0(v) \leq N(v) \leq MN_0(v) \quad \forall v \in \mathbf{V}$$

Per $v = \underline{0}$ la tesi e' vera. Consideriamo quindi solo vettori non nulli.

(2.1) Dim. $\exists M$

Sia $H = \max_{i=1,\dots,n} N(e_i)$

$$N(v) = N\left(\sum_{i=1}^n v_i e_i\right) \leq \sum_{i=1}^n N(v_i e_i) = \sum_{i=1}^n |v_i| N(e_i) \leq \sum_{i=1}^n N_0(v) H = n H N_0(v)$$

Quindi $M = nH$

(2.2) Dim $\exists m$

$(\mathbf{V}, d)$ e' uno spazio metrico, dove

$$d(v, w) = N_0(w - v)$$

e' una metrica.

N e' allora una funzione dallo spazio metrico $(\mathbf{V}, d)$ allo spazio metrico $(\mathbb{R}, d_e)$, dove d_e e' la norma euclidea. Possiamo quindi applicare a N i thm di analisi II.

Si prova che $N_0()$ e' una funzione continua in $\mathbf{V}$.

(2.3) $N()$ e' continua

$$N_0 \text{ continua in } x \Rightarrow \lim_{x \rightarrow x_0} N_0(x) = N_0(x_0) \Leftrightarrow \lim_{x \rightarrow x_0} |N_0(x) - N_0(x_0)| = 0$$

$$\left\{ \begin{array}{l} 0 \leq |N(x) - N(x_0)| \underset{2.1}{\leq} M |N_0(x) - N_0(x_0)| \\ 0 \longrightarrow 0 \\ M |N_0(x_n) - N_0(x)| \longrightarrow 0 \end{array} \right. \underset{[AII, 8.27, \text{pg. 85}]}{\Rightarrow} \lim_{x \rightarrow x_0} |N(x) - N(x_0)| = 0 \Leftrightarrow \lim_{x \rightarrow x_0} N(x) = N(x_0)$$

Quindi $N()$ e' continua.

(2.4) $\exists m$

Sia data la n -sfera secondo la norma N_0 :

$$S_n = \{v \in \mathbf{V} \mid N_0(v) = 1\}$$

S_n e' un insieme sequenzialmente compatto.

Allora, poiche' N e' continua in S_n , per il thm di Weierstrass (vedi [AII, 8.34, pg. 88]) si ha

$$\exists c = \min N(S_n), \quad \exists C = \max N(S_n) > 0$$

allora,

$$\frac{v}{N_0(v)} \in S_n \Rightarrow N\left(\frac{v}{N_0(v)}\right) = \frac{N(v)}{N_0(v)} \geq c \Leftrightarrow N(v) \geq c N_0(v)$$

Quindi $m = c$

□

Definition 4.6. Sia $A \in \mathbb{K}^{n \times m}$, utilizzando le norme viste in [4.1, pg. 4], definiamo

$$\|A\|_{p,q} \stackrel{\text{def}}{=} \sup \left\{ \frac{\|Ax\|_p}{\|x\|_q} \mid x \in \mathbb{K}^{m \times 1}, x \neq 0 \right\}$$

$$\|A\|_p \stackrel{\text{def}}{=} \|A\|_{p,p}$$

chiamata norma matriciale, indotta da $\|\cdot\|_p, \|\cdot\|_q$.

Definiamo la trasposta coniugata di una matrice complessa:

$${}^{\tau}A = \overline{{}^tA}$$

ovvero, ${}^{\tau}A$, e' la matrice coniugata² della trasposta di A . Se A e' reale, la trasposta coniugata coincide con l'usuale trasposta.

² $(\overline{A})_{ij} = \overline{a_{ji}}$

Useremo anche la seguente convenzione:

$$|a| = \sqrt{a\bar{a}} = \begin{cases} \text{valore assoluto di } a & a \in \mathbb{R} \\ \text{modulo di } a & a \in \mathbb{C} \end{cases}$$

Definiamo il *raggio spettrale* di A :

$$\rho(A) = \max\{|\lambda| \mid \lambda \text{ autovalore di } A\}$$

Il prodotto scalare su $\mathbb{K}^n$, con $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$, e' definito come:

$$\begin{aligned} \text{se } x, y \in \mathbb{K}^{1,n} : \quad x \cdot y &\stackrel{\text{def}}{=} x \tau y \\ \text{se } x, y \in \mathbb{K}^{n,1} : \quad x \cdot y &\stackrel{\text{def}}{=} \tau xy \end{aligned}$$

Se $\mathbb{K} = \mathbb{R}$, $\cdot$ si riduce al prodotto scalare standard.

Proposition 4.7. *Sia $A \in \mathbb{K}^{n \times m}$, si ha*

$$1. \quad \|A\|_{p,q} = \max \left\{ \|Ax\|_p \mid x \in \mathbb{K}^{m \times 1}, \|x\|_q = 1 \right\}$$

Supponendo $\mathbb{K} = \mathbb{R}$, si ha

$$2. \quad \|A\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{ij}|$$

$$3. \quad \|A\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^n |a_{ij}|$$

$$4. \quad \|A\|_1 = \|^t A\|_{\infty}$$

$$5. \quad \|A\|_2 = \sqrt{\rho(^t A A)}$$

$$5.1 \quad A = ^t A \Rightarrow \|A\|_2 = \rho(A)$$

dove ρ e' il raggio spettrale.

Proof:

(1) Dim 1.

Sia

$$\lambda = \|x\|_q$$

$$\bar{x} = \frac{x}{\|x\|_q} \quad \text{e' il versore di } x$$

$$\|\bar{x}\|_q = 1 \quad (*)$$

$$x = \lambda \bar{x}$$

Osserviamo che:

$$\frac{\|Ax\|_p}{\|x\|_q} = \frac{\|A\lambda\bar{x}\|_p}{\|\lambda\bar{x}\|_q} = \frac{|\lambda|\|A\bar{x}\|_p}{|\lambda|\|\bar{x}\|_q} = \frac{\|A\bar{x}\|_p}{\|\bar{x}\|_q} \underbrace{=}_{(*)} \|A\bar{x}\|_p$$

Quindi, basta limitarsi a considerare i versori, ovvero consideriamo tutti e soli i vettori di $S_n = \left\{ x \mid x \in \mathbb{K}^{m \times 1}, \|x\|_q = 1 \right\}$

Consideriamo la funzione $N'(x) = \|Ax\|_p$. $N'(x)$ e' continua, perche' composizione di funzioni continue³. Allora, procedendo come nella dim. del thm [4.5,pg.5], deduciamo che $N'(x)$ ammette minimo e massimo su S_n , e quindi

$$\sup_{x \in S_n} N'(x) = \max_{x \in S_n} N'(x)$$

(2) Dim. 2.

Ridefiniamo S_n secondo la norma $\| - \|_\infty$:

$$S_n = \{x \mid x \in \mathbb{K}^{m \times 1}, \|x\|_\infty = 1\}$$

Si ha

$$\|A\|_\infty = \underbrace{\max}_{1.} \|Ax\|_\infty \quad (2)$$

$$\|Ax\|_\infty = \max_{1 \leq i \leq n} |(Ax)_i| = \max_{1 \leq i \leq n} \left| \sum_{j=1}^m a_{ij} x_j \right| \quad (2.1)$$

(2.1) $\exists c \geq 0 : \|Ax\|_\infty \leq c \forall x \in S_n$

$$x \in S_n \Rightarrow \|x\|_\infty = 1 \Leftrightarrow \max |x_h| = 1 \Leftrightarrow |x_h| \leq 1 \forall h = 1, 2, \dots, m \quad (*)$$

$$\left| \sum_{j=1}^m a_{ij} x_j \right| \leq \sum_{j=1}^m |a_{ij}| |x_j| \leq \sum_{j=1}^m |a_{ij}| =: c_i$$

$$\text{Sia } c = \max_{i=1, \dots, n} c_i$$

$$\|Ax\|_\infty \stackrel{(2.1)}{\leq} \max_{1 \leq i \leq n} c_i = c$$

(2.2) $\exists \bar{x} \in S_n : \|A\bar{x}\|_\infty \geq c$

$$\begin{aligned} c &= \max_{i=1, \dots, n} c_i = c_k = \sum_{j=1}^m |a_{kj}| \\ &\Rightarrow c \geq \sum_{j=1}^m |a_{ij}| \quad \forall i = 1, \dots, n \quad (3) \end{aligned}$$

Scegliamo la seguente x :

$$x_j = \begin{cases} 1 & a_{kj} \geq 0 \\ -1 & a_{kj} < 0 \end{cases} \quad (**)$$

si ha

$$(Ax)_k = \sum_{j=1}^m a_{kj} x_j \stackrel{(**)}{=} \sum_{j=1}^m |a_{kj}| = c_k = c$$

$$\|Ax\|_\infty = \max_{1 \leq i \leq n} |(Ax)_i| \geq |(Ax)_k| = c$$

$$\|Ax\|_\infty \geq c$$

(2.3) Q.E.D.

Sia

$$\|Ay\|_\infty = \max_{x \in S_n} \|Ax\|_\infty$$

Il passo 2.1 implica che

$$\forall x \in S_n \quad \|Ax\|_\infty \leq c \Rightarrow \|Ay\|_\infty \leq c \Leftrightarrow \max_{x \in S_n} \|Ax\|_\infty \leq c$$

³ Ax e' continua perche' composizione di funzioni lineari. $N(x) = \|x\|_p$ e' continua per quanto visto in [4.5,pg.5]

Il passo 2.2 implica che

$$\max_{x \in S_n} \|Ax\|_\infty \geq \|A\bar{x}\|_\infty \geq c$$

Quindi

$$\max_{x \in S_n} \|Ax\|_\infty \leq c \Leftrightarrow \max_{x \in S_n} \|Ax\|_\infty = c \Leftrightarrow \|A\|_\infty = c = \max_{i=1, \dots, n} \sum_{j=1}^m |a_{ij}|$$

(3) Dim 4.

Immediata conseguenza della 2., 3..

(4) Dim 5.

(4.1) Osserviamo che dati n numeri reali o complessi $s_1, \dots, s_n$, si ha

$$(\max\{|s_1|, \dots, |s_n|\})^2 = \max\{|s_1|^2, \dots, |s_n|^2\}$$

Infatti, sia $|s| = \max\{|s_1|, \dots, |s_n|\}$, allora

$$|s| \geq |s_i| \quad \forall i = 1, \dots, n \Rightarrow |s|^2 \geq |s_i|^2 \quad \forall i = 1, \dots, n \Rightarrow |s|^2 = \max\{|s_1|^2, \dots, |s_n|^2\}$$

Abbiamo:

$$\|A\|_2 \stackrel{1.}{=} \max\{\|Ax\|_2 \mid \|x\|_2 = 1\}$$

(4.2) Dim. che $\|Ax\|_2^2 = {}^t x C x$, con $C \in \text{Sym}_m(\mathbb{R})$

Fissiamo adesso una $x : \|x\|_2 = 1$, si ha

$$\|Ax\|_2 = \sqrt{\sum_{i=1}^n |(Ax)_i|^2}$$

$$\|Ax\|_2^2 = Ax \cdot Ax = {}^t(Ax)Ax = {}^t x A A x$$

Sia $C = {}^t A A \in \mathbb{R}^{m \times m}$. C e' simmetrica, infatti,

$${}^t C = {}^t ({}^t A A) = {}^t A A = C$$

Allora per quanto visto in Geometria II:

$C \in \text{Sym}_m(\mathbb{R}) \stackrel{[GII, 4.13, pg. 72]}{\Rightarrow} \exists B = \{u_1, \dots, u_m\}$ base ortonormale di $\mathbb{R}^m$: u_i sono autovettori di C

$$B \text{ ortonormale} \Rightarrow \begin{cases} \|u_i\|_2 = 1 \quad \forall i = 1, \dots, m \\ u_i \perp u_j \Leftrightarrow u_i \cdot u_j = 0 \quad \forall i \neq j \end{cases}$$

Sia λ_i l'autovalore di u_i , ovvero

$$C u_i = \lambda_i u_i$$

Per il lemma [GII, 4.11, pg. 70], si ha $\lambda_i \in \mathbb{R} \quad \forall i = 1, \dots, m$.

(4.3) Dim. che $\max\{|\lambda_i| \mid i = 1, \dots, m\} = \rho(C)$

Se costruiamo B usando il metodo descritto in [GII, 4.12, pg. 71], allora $\{\lambda_i \mid i = 1, \dots, m\}$ e'

l'insieme di tutti gli autovalori di C (puo' anche darsi che $\lambda_i = \lambda_j$ per qualche $i \neq j$).

Quindi, per definizione (vedi [4.6, pg. 7]), si ha

$$\max\{|\lambda_i| \mid i = 1, \dots, m\} = \rho(C)$$

(4.4) $\lambda_i \geq 0 \quad \forall i = 1, \dots, m$

$$\|A u_i\|_2^2 \stackrel{\text{passo 4.2}}{=} {}^t u_i C u_i = {}^t u_i \lambda_i u_i = \lambda_i {}^t u_i u_i = \lambda_i u_i \cdot u_i = \lambda_i \|u_i\|_2^2$$

$$\lambda_i \|u_i\|_2^2 = \|A u_i\|_2^2 \geq 0 \Leftrightarrow \lambda_i \geq 0$$

⟨4.5⟩ Dim. che

$$\|Ax\|_2^2 = \sum_{i=1}^m |\lambda_i| |\alpha_i|^2$$

$$\|x\|_2^2 = \sum_{i=1}^m |\alpha_i|^2$$

dove le α_i sono le componenti di x secondo la base B

$$\begin{aligned} \left\{ \begin{array}{l} B \text{ base ortonormale di } \mathbb{R}^m \\ x \in \mathbb{R}^m \end{array} \right. &\Rightarrow x = \sum_{i=1}^m \alpha_i u_i \\ \|Ax\|_2^2 &\stackrel{\text{passo 4.2}}{=} {}^t x C x = \sum_{i=1}^m {}^t x C \alpha_i u_i = \sum_{i=1}^m \sum_{j=1}^m {}^t (\alpha_j u_j) C \alpha_i u_i = \sum_{i=1}^m \sum_{j=1}^m \alpha_j \alpha_i {}^t u_j C u_i \stackrel{C u_i = \lambda_i u_i}{=} \sum_{i=1}^m \sum_{j=1}^m \alpha_j \alpha_i \lambda_i {}^t u_j u_i = \\ &= \sum_{i=1}^m \sum_{j=1}^m \alpha_j \alpha_i \lambda_i u_j \cdot u_i \stackrel{u_i \cdot u_j = 0 \quad \forall i \neq j}{=} \sum_{i=1}^m |\alpha_i|^2 \lambda_i u_i \cdot u_i = \sum_{i=1}^m \lambda_i |\alpha_i|^2 \|u_i\|_2^2 \stackrel{\|u_i\|_2 = 1}{=} \sum_{i=1}^m \lambda_i |\alpha_i|^2 \stackrel{\text{passo 4.4}}{=} \sum_{i=1}^m |\lambda_i| |\alpha_i|^2 \end{aligned}$$

Infine,

$$\begin{aligned} B \text{ ortonormale} &\stackrel{[GII, 3.13, \text{pg.59}]}{\Rightarrow} x \cdot x = {}^t [x]_B [x]_B = \sum_{i=1}^m |\alpha_i|^2 \\ \|x\|_2^2 &= x \cdot x = \sum_{i=1}^m |\alpha_i|^2 \end{aligned}$$

⟨4.6⟩ Q.E.D.

$$\begin{aligned} \|A\|_2^2 &= (\max \{ \|Ax\|_2 \mid \|x\|_2 = 1 \})^2 \stackrel{\text{passo 4.1}}{=} \max \{ \|Ax\|_2^2 \mid \|x\|_2 = 1 \} \\ &\stackrel{\text{passo 4.5}}{=} \max \left\{ \sum_{i=1}^m |\lambda_i| |\alpha_i|^2 \mid \sum_{i=1}^m |\alpha_i|^2 = 1 \right\} \quad (**) \\ \left\{ \begin{array}{l} |\lambda_i| \geq 0 \\ \sum_{i=1}^m |\alpha_i|^2 = 1 \end{array} \right. &\Rightarrow \sum_{i=1}^m |\lambda_i| |\alpha_i|^2 \text{ e' una media pesata il cui max si ha per } \alpha_i = \delta_{ki} \quad \forall i = 1, \dots, m \\ &\text{dove } k \text{ e' tale che } |\lambda_k| = \max \{ |\lambda_i| \}_i \text{ Ovvero, si ha:} \\ \max \left\{ \sum_{i=1}^m |\lambda_i| |\alpha_i|^2 \mid \sum_{i=1}^m |\alpha_i|^2 = 1 \right\} &= |\lambda_k| \stackrel{\text{passo 4.3}}{=} \rho(C) \\ &\stackrel{(**)}{\Rightarrow} \|A\|_2^2 = \rho(C) \Leftrightarrow \|A\|_2 = \sqrt{\rho({}^t A A)} \end{aligned}$$

Infine, se A e' simmetrica, si ha:

$$A \text{ simmetrica} \Leftrightarrow A = {}^t A \Rightarrow \rho(C) = \rho({}^t A A) = \rho(AA)$$

⟨4.6.1⟩ Siano $L = \{\lambda_1, \dots, \lambda_m\}$. $M = \{\mu_1^2, \dots, \mu_m^2\}$ rispettivamente l'insieme di tutti gli autovalori di $C = AA$ e di tutti i quadrati degli autovalori di A .

Si ha che $\mu_1, \dots, \mu_m$ sono tutti gli autovalori di A , e inoltre $L = M$.

Poiche' A e' simmetrica,

$$\begin{aligned} A \in \text{Sym}_m(\mathbb{R}) & \xRightarrow{\text{[GII,4.13,pg.72]}} \begin{cases} \exists \{v_1, \dots, v_m\} \text{ base ortonormale di } \mathbb{R}^m, \\ v_1, \dots, v_m \text{ autovettori di } A: Av_i = \mu_i v_i \\ \{\mu_1, \dots, \mu_m\} \text{ insieme di tutti gli autovettori di } A \end{cases} \\ \langle 4.6.1.1 \rangle \text{ Dim } L \supseteq M & \end{aligned}$$

$$\begin{aligned} \mu_i \in M & \Rightarrow Av_i = \mu_i v_i \\ Cv_i &= AA v_i = A \mu_i v_i = \mu_i^2 v_i \\ Cv_i &= \mu_i^2 v_i \Rightarrow \mu_i^2 \in L \end{aligned}$$

$$\langle 4.6.1.2 \rangle \text{ Dim } L \subseteq M$$

$$\lambda_i \in L \Rightarrow Cu_i = \lambda_i u_i \quad [*]$$

$$u_i = \sum_{j=1}^m \beta_j v_j$$

$$u_i \in B \text{ base} \Rightarrow u_i \neq 0 \Rightarrow \exists \beta_k \neq 0$$

$$Cu_i = AA u_i = \sum_{j=1}^m AA \beta_j v_j = \sum_{j=1}^m \beta_j \mu_j^2 v_j \underset{[*]}{=} \lambda_i u_i = \lambda_i \sum_{j=1}^m \beta_j v_j$$

$$\sum_{j=1}^m \beta_j \mu_j^2 v_j = \sum_{j=1}^m \lambda_i \beta_j v_j \Leftrightarrow \sum_{j=1}^m (\beta_j \mu_j^2 - \lambda_i \beta_j) v_j = 0 \quad \Leftrightarrow \quad \beta_j \mu_j^2 - \lambda_i \beta_j = 0 \quad \forall j = 1, \dots, m \Rightarrow$$

$v_1, \dots, v_m \text{ sono l.i.}$

$$\Rightarrow \beta_k \mu_k^2 = \lambda_i \beta_k \underset{k \neq 0}{\Rightarrow} \lambda_i = \mu_k^2$$

$$\lambda_i = \mu_k^2 \Rightarrow \lambda_i \in M$$

$$\langle 4.6.2 \rangle \rho(AA) = \rho(A)^2$$

$$A \in \text{Sym}_m(\mathbb{R}) \xRightarrow{\text{[GII,4.11,pg.70]}} \mu_1, \dots, \mu_m \in \mathbb{R} \Rightarrow |\mu_i|^2 = \mu_i^2 \quad \forall i \quad [**]$$

$$\rho(AA) = \max_{\text{passo 4.6.1}} L \underset{\text{passo 4.6.1}}{=} \max M = \max \{\mu_1^2, \dots, \mu_m^2\} \underset{[**]}{=} \max \{|\mu_1|^2, \dots, |\mu_m|^2\} =$$

$$\underset{\text{passo 4.1}}{=} (\max \{|\mu_1|, \dots, |\mu_m|\})^2 = \rho(A)^2$$

$$\rho(AA) = \rho(A)^2$$

In conclusione,

$$\|A\|_2 = \sqrt{\rho(AA)} \underset{\text{passo 4.6.2}}{=} \sqrt{\rho(A)^2} = |\rho(A)| = \rho(A)$$

□

Proposition 4.8. La funzione $\| - \|_{p,q} : \mathbb{K}^{n \times m} \longrightarrow \mathbb{R}_+$ e' una norma. Inoltre, valgono le seguenti proprieta':

$$1. \|Ax\|_p \leq \|A\|_{p,q} \|x\|_q \quad \forall x \in \mathbb{K}^{m,1}$$

Inoltre, data $B \in \mathbb{R}^{m \times r}$ e data la norma $\| - \|_{q,q} : \mathbb{K}^{m \times r} \longrightarrow \mathbb{R}_+$, si ha

$$2. \|AB\|_{p,q} \leq \|A\|_{p,q} \|B\|_{q,q}$$

Proof:

(1) $\|-\|_{p,q}$ e' una norma

1. $\|A\|_{p,q} \geq 0$, infatti,

$$\|A\|_{p,q} \underbrace{=}_{[4.7, \text{pg. 7}]} \max \left\{ \|Ax\|_p \mid \|x\|_q = 1 \right\} = \|A\bar{x}\|_p \underbrace{\geq}_{\|-\|_p \text{ e' una norma}} 0$$

2. $\|A\|_{p,q} = 0 \Leftrightarrow A = \underline{0}^{n \times m}$, infatti,

$$\|A\|_{p,q} = \max \left\{ \frac{\|Ax\|_p}{\|x\|_q} \mid x \neq 0 \right\} \Leftrightarrow \|Ax\|_p = 0 \quad \forall x \neq 0 \quad \underbrace{\Leftrightarrow}_{\|-\|_p \text{ e' una norma}} Ax = 0 \quad \forall x \neq 0$$

Allora, scegliendo x in modo opportuno, si ha:

Per $x_i = \delta_{1i}$, $i = 1, \dots, m$ si ha $Ax = A^1$, $Ax_p = 0 \Leftrightarrow A^1 = 0$

Per $x_i = \delta_{2i}$, $i = 1, \dots, m$ si ha $Ax = A^2$, $Ax_p = 0 \Leftrightarrow A^2 = 0$

$\vdots$

Per $x_i = \delta_{ni}$, $i = 1, \dots, m$ si ha $Ax = A^n$, $Ax_p = 0 \Leftrightarrow A^n = 0$

Quindi tutte le colonne di A sono nulle, ovvero A e' la matrice nulla.

3. $\|A+B\|_{p,q} \leq \|A\|_{p,q} + \|B\|_{p,q}$, infatti,

$$\begin{aligned} \|A+B\|_{p,q} &= \max \left\{ \|(A+B)x\|_p = \|Ax+Bx\|_p \mid \|x\|_q = 1 \right\} = \|A\bar{x} + B\bar{x}\|_p \leq \\ &\leq \|A\bar{x}\|_p + \|B\bar{x}\|_p \leq \|A\|_{p,q} + \|B\|_{p,q} \end{aligned}$$

(2) Dim 1.

Se $\|x\|_q = 0 \Leftrightarrow x = 0$, la tesi e' vera. Consideriamo quindi $x \neq 0$.

$$\begin{aligned} \|A\|_{p,q} &\underbrace{=}_{[4.7, \text{pg. 7}]} \max \left\{ \frac{\|Ax\|_p}{\|x\|_q} \mid x \in \mathbb{K}^{m \times 1}, x \neq 0 \right\} \geq \frac{\|Ax\|_p}{\|x\|_q} \quad \forall x \in \mathbb{K}^{m \times 1}, x \neq 0 \Leftrightarrow \\ &\Leftrightarrow \|A\|_{p,q} \|x\|_q \geq \|Ax\|_p \quad \forall x \in \mathbb{K}^{m \times 1}, x \neq 0 \end{aligned}$$

(3) Dim 2.

$$x \in \mathbb{R}^{r \times 1}, B \in \mathbb{R}^{m \times r} \Rightarrow Bx \in \mathbb{R}^{m \times 1} \quad (1)$$

$$\|ABx\|_p = \|A(Bx)\|_p \leq \|A\|_{p,q} \|Bx\|_q \underbrace{\leq}_{1.} \|A\|_{p,q} \|B\|_{q,q} \|x\|_q$$

$$\frac{\|ABx\|_p}{\|x\|_q} \leq \|A\|_{p,q} \|B\|_{q,q} \quad \forall x \in \mathbb{R}^{m \times 1} \Rightarrow \sup \left\{ \frac{\|ABx\|_p}{\|x\|_q} \right\} \leq \|A\|_{p,q} \|B\|_{q,q} \underbrace{\Leftrightarrow}_{[4.6, \text{pg. 6}]} \|AB\|_{p,q} \leq \|A\|_{p,q} \|B\|_{q,q}$$

□

Proposition 4.9. Sia $A \in \mathbb{R}^{n \times n}$, si ha

1. $\rho(A) \leq \|A\|_{p,q} \quad \forall p, q \in \mathbb{N}^*$

2. $\forall A \in \mathbb{R}^{n \times n} \quad \forall \varepsilon > 0 \quad \exists p, q \in \mathbb{R}^*, \quad \text{dipendenti da } A \text{ e da } \varepsilon, \text{ tali che } \left| \|A\|_{p,q} - \rho(A) \right| < \varepsilon$

3. $\rho(A) = \inf_{p, q \in \mathbb{N}^*} \|A\|_{p,q}$

Proposition 4.10. Sia $A \in \mathbb{R}^{n \times n}$ invertibile. Sia $\lambda_m({}^tAA)$ l'autovalore minimo in modulo di

${}^tAA$ e $\lambda_M({}^tAA)$ l'autovalore massimo. Si ha

$$\begin{aligned} 1. \quad \|A\|_2 &= \sqrt{\lambda_M({}^tAA)} \\ 2. \quad \|A^{-1}\|_2 &= \frac{1}{\sqrt{\lambda_m({}^tAA)}} \end{aligned}$$

Proof: La 1. si ha direttamente per la [4.7,pg.7].

Dimostriamo la 2.

Notiamo intanto che se λ e' un autovalore di B allora $\frac{1}{\lambda}$ e' un autovalore di B^{-1} , infatti,

$$Bu = \lambda u \Rightarrow u = B^{-1}\lambda u$$

$$\frac{1}{\lambda}u = B^{-1}u$$

Osserviamo che ${}^tAA$ e A^tA hanno gli stessi autovalori, infatti,

$$A({}^tAA)A^{-1} = A^tA \Rightarrow {}^tAA, A^tA \text{ sono simili } \Rightarrow \text{hanno gli stessi autoval.} \quad (1)$$

[AL,8.7,pg.53]

Se con $l_1, \dots, l_n$ indichiamo gli autovalori di ${}^tAA$, e li ordiniamo in modo da avere

$$\lambda_M({}^tAA) = |l_1| \geq |l_2| \geq \dots \geq |l_n| = \lambda_m({}^tAA)$$

elevando i singoli membri per -1 , abbiamo

$$\frac{1}{|l_1|} \leq \frac{1}{|l_2|} \leq \dots \leq \frac{1}{|l_n|}$$

che sono gli autovalori di $A^{-1}{}^tA^{-1}$, e quindi

$$\rho(A^{-1}) \underbrace{=}_{[4.7,pg.7]} \sqrt{\rho({}^tA^{-1}A^{-1})} = \sqrt{\lambda_M({}^tA^{-1}A^{-1})} = \sqrt{\lambda_M((A^tA)^{-1})} = \frac{1}{\sqrt{\lambda_m(A^tA)}} \underbrace{=}_{(1)} \frac{1}{\sqrt{\lambda_m({}^tAA)}}$$

□

Proposition 4.11. Sia $A \in \mathbb{R}^{n \times n}$ invertibile e simmetrica. Siano $\{\lambda_i\}$ i suoi autovalori, ordinati come segue:

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$$

Si ha

$$\begin{aligned} \|A\|_2 &= |\lambda_1| \\ \|A^{-1}\|_2 &= \frac{1}{|\lambda_n|} \end{aligned}$$

Proof: Basta ripetere la dimostrazione della prop [4.10,pg.12], considerando che nel caso di A simmetrica si ha

$$\|A\|_2 = \rho(A) = \lambda_M(A)$$

e che anche A^{-1} e' simmetrica, infatti,

$${}^t(A^{-1}) = ({}^tA)^{-1} = A^{-1}$$

□

5 Sistemi lineari

Proposition 5.1. Sia dato un sistema triangolare superiore

$$Ax = b, \text{ con } A \in \mathbb{R}^{n \times n}, |A| \neq 0, x, b \in \mathbb{R}^n$$

Si ha:

$$x_i = \frac{b_i - \sum_{j=i+1}^n a_{ij}x_j}{a_{ii}} \quad \forall i = n, n-1, \dots, 1$$

Se $Ax = b$ e' un sistema triangolare inferiore, si ha:

$$x_i = \frac{b_i - \sum_{j=1}^{i-1} a_{ij}x_j}{a_{ii}} \quad \forall i = 1, \dots, n$$

Il costo di questa soluzione e':

$$\frac{n(n+1)}{2}$$

Proof:

$$\begin{aligned} |A| &\stackrel{\text{A matrice triangolare}}{=} \prod_{i=1}^n a_{ii} \quad (*) \\ |A| \neq 0 &\stackrel{(*)}{\Rightarrow} a_{ii} \neq 0 \quad \forall i = 1, \dots, n \end{aligned}$$

(1) Triangolare superiore

$$\begin{aligned} Ax = b &\stackrel{\text{A mat. triang. sup.}}{\Leftrightarrow} \Leftrightarrow \sum_{j=i}^n a_{ij}x_j = b_i \quad \forall i = n, n-1, \dots, 1 \Leftrightarrow \\ &\Leftrightarrow a_{ii}x_i + \sum_{j=i+1}^n a_{ij}x_j = b_i \quad \forall i = n, n-1, \dots, 1 \Leftrightarrow \\ &\Leftrightarrow x_i = \frac{b_i - \sum_{j=i+1}^n a_{ij}x_j}{a_{ii}} \quad \forall i = n, n-1, \dots, 1 \end{aligned}$$

(2) Triangolare inferiore

$$\begin{aligned} Ax = b &\stackrel{\text{A mat. triang. inf.}}{\Leftrightarrow} \Leftrightarrow \sum_{j=1}^i a_{ij}x_j = b_i \quad \forall i = 1, \dots, n \Leftrightarrow \\ &\Leftrightarrow a_{ii}x_i + \sum_{j=1}^{i-1} a_{ij}x_j = b_i \quad \forall i = 1, \dots, n \Leftrightarrow \\ &\Leftrightarrow x_i = \frac{b_i - \sum_{j=1}^{i-1} a_{ij}x_j}{a_{ii}} \quad \forall i = 1, \dots, n \end{aligned}$$

(3) Calcoliamo il costo

$$\begin{aligned} [x_i] &= \left[\left(\sum_{j=i+1}^n a_{ij} * x_j \right) / a_{ii} \right] = \left[\sum_{j=i+1}^n a_{ij} * x_j \right] + [/ a_{ii}] = \\ &= \sum_{j=i+1}^n [a_{ij} * x_j] + 1 = (n - (i+1) + 1) + 1 = n - i + 1 \\ [x_i \quad \forall i = n, n-1, \dots, 1] &= \sum_{i=1}^n [x_i] = \sum_{i=1}^n n - i + 1 = n^2 + n - \sum_{i=1}^n i \stackrel{[2.1, pg.2]}{=} n^2 + n - \frac{n(n+1)}{2} = \frac{n(n+1)}{2} \end{aligned}$$

□

Proposition 5.2. *Riduzione di un sistema lineare a un sistema triangolare (riduzione naive di Gauss).*

Sia dato un sistema

$$Ax = b, \text{ con } A \in \mathbb{R}^{n \times n}, |A| \neq 0, x, b \in \mathbb{R}^n$$

Il seguente e' un algoritmo che riduce la matrice completa $A^1 = (A|b)$ del sistema $Ax = b$ alla matrice completa A^n di un sistema triangolare equivalente ad $Ax = b$:

$$\left(A_i^{k+1} := A_i^k - \frac{a_{ik}^k}{a_{kk}^k} A_k^k \quad \forall i = k+1, \dots, n \right) \quad \forall k = 1, \dots, n-1$$

dove A_i^k rappresenta l' i -esima riga, della matrice A^k .

Attenzione: stiamo supponendo che $a_{kk}^k \neq 0 \quad \forall k = 1, \dots, n-1$.

In forma esplicita, l'algoritmo e':

$$\left(\left(a_{ij}^{k+1} := a_{ij}^k - \frac{a_{ik}^k}{a_{kk}^k} a_{kj}^k \quad \forall j = k, \dots, n+1 \right) \quad \forall i = k+1, \dots, n \right) \quad \forall k = 1, \dots, n-1$$

Il codice ottimizzato:

$$\begin{aligned} & \forall k = 1, \dots, n-1 \\ & \quad \forall i = k+1, \dots, n \\ & \quad \quad m := \frac{a_{ik}^k}{a_{kk}^k} \\ & \quad \quad \forall j = k, \dots, n+1 \\ & \quad \quad \quad a_{ij}^{k+1} := a_{ij}^k - m a_{kj}^k \end{aligned}$$

Il costo di quest'ultimo codice e'

$$3 \frac{(n-1)n}{2} + \frac{2(n-1)^3 + 3(n-1)^2 + n-1}{6} \approx \frac{n^3}{3} + O(n^2)$$

Il numero $m_{ik} = \frac{a_{ik}^k}{a_{kk}^k}$, viene chiamato moltiplicatore di posto i, k .

Proof: PROOF SKETCH: Sia $A^1 = (A|b)$ la matrice completa del sistema $Ax = b$. Ridurremo il sistema iniziale A^1 , nel sistema equivalente A^2 , poi A^2 in A^3 , etc...

L'idea di base consiste nell'effettuare riduzioni del seguente tipo:

$$\begin{aligned}
A^1 &= \left[\begin{array}{cccc|c} a_{11}^1 & a_{12}^1 & a_{13}^1 & \dots & a_{1n}^1 \\ a_{21}^1 & a_{22}^1 & a_{23}^1 & \dots & a_{2n}^1 \\ a_{31}^1 & a_{32}^1 & a_{33}^1 & \dots & a_{3n}^1 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1}^1 & a_{n2}^1 & a_{n3}^1 & \dots & a_{nn}^1 \end{array} \right] a_{1n+1}^1 \longrightarrow \\
&\longrightarrow A^2 = \left[\begin{array}{cccc|c} a_{11}^2 & a_{12}^2 & a_{13}^2 & \dots & a_{1n}^2 \\ 0 & a_{22}^2 & a_{23}^2 & \dots & a_{2n}^2 \\ 0 & a_{32}^2 & a_{33}^2 & \dots & a_{3n}^2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & a_{n2}^2 & a_{n3}^2 & \dots & a_{nn}^2 \end{array} \right] a_{1n+1}^2 \longrightarrow \\
&\longrightarrow A^3 = \left[\begin{array}{cccc|c} a_{11}^3 & a_{12}^3 & a_{13}^3 & \dots & a_{1n}^3 \\ 0 & a_{22}^3 & a_{23}^3 & \dots & a_{2n}^3 \\ 0 & 0 & a_{33}^3 & \dots & a_{3n}^3 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & a_{n3}^3 & \dots & a_{nn}^3 \end{array} \right] a_{1n+1}^3 \longrightarrow \dots \longrightarrow A^n
\end{aligned}$$

Procediamo con la dimostrazione: sia A_i^1 la riga i -esima di A^1 . Eseguiamo le seguenti operazioni elementari⁴ su A^1 :

$$\begin{aligned}
A_2^1 &\rightarrow A_2^1 - \frac{a_{21}^1}{a_{11}^1} A_1^1 \\
A_3^1 &\rightarrow A_3^1 - \frac{a_{31}^1}{a_{11}^1} A_1^1 \\
&\vdots \\
A_n^1 &\rightarrow A_n^1 - \frac{a_{n1}^1}{a_{11}^1} A_1^1
\end{aligned}$$

ottieniamo così A^2 .
Adesso lavoriamo su A^2 :

$$\begin{aligned}
A_3^2 &\rightarrow A_3^2 - \frac{a_{32}^2}{a_{22}^2} A_2^2 \\
A_4^2 &\rightarrow A_4^2 - \frac{a_{42}^2}{a_{22}^2} A_2^2 \\
&\vdots \\
A_n^2 &\rightarrow A_n^2 - \frac{a_{n2}^2}{a_{22}^2} A_2^2
\end{aligned}$$

⁴le operazioni elementari che non cambiano il sistema sono: scambio di righe, moltiplicazione di una riga per $\lambda \neq 0$, aggiungere a una riga una combinazione lineare di altre righe

ottieniamo cosi' $A^{\mathbf{3}}$.

In generale, per ricavare $A^{\mathbf{k}+1}$:

$$\begin{aligned} A_{k+1}^{\mathbf{k}} &\rightarrow A_{k+1}^{\mathbf{k}} - \frac{a_{k+1k}^{\mathbf{k}}}{a_{kk}^{\mathbf{k}}} A_k^{\mathbf{k}} \\ A_{k+2}^{\mathbf{k}} &\rightarrow A_{k+2}^{\mathbf{k}} - \frac{a_{k+2k}^{\mathbf{k}}}{a_{kk}^{\mathbf{k}}} A_k^{\mathbf{k}} \\ &\vdots \\ A_n^{\mathbf{k}} &\rightarrow A_n^{\mathbf{k}} - \frac{a_{nk}^{\mathbf{k}}}{a_{kk}^{\mathbf{k}}} A_k^{\mathbf{k}} \end{aligned}$$

E in forma piu' compatta:

$$A_i^{\mathbf{k}+1} := A_i^{\mathbf{k}} - \frac{a_{ik}^{\mathbf{k}}}{a_{kk}^{\mathbf{k}}} A_k^{\mathbf{k}} \quad \forall i = k+1, \dots, n$$

In definitiva, poiche' dobbiamo arrivare fino alla matrice completa $A^{\mathbf{n}}$, ovvero con $k+1 = n \Leftrightarrow k = n-1$, l'algoritmo completo e':

$$\left(A_i^{\mathbf{k}+1} := A_i^{\mathbf{k}} - \frac{a_{ik}^{\mathbf{k}}}{a_{kk}^{\mathbf{k}}} A_k^{\mathbf{k}} \quad \forall i = k+1, \dots, n \right) \quad \forall k = 1, \dots, n-1$$

Poiche' nella riga $A_k^{\mathbf{k}}$ tutti gli elementi $a_{kj}^{\mathbf{k}}$, $j < k$ sono gia' nulli, possiamo evitare le moltiplicazioni nulle, scrivendo direttamente le componenti di $A_i^{\mathbf{k}+1}$:

$$\left(\left(a_{ij}^{\mathbf{k}+1} := a_{ij}^{\mathbf{k}} - \frac{a_{ik}^{\mathbf{k}}}{a_{kk}^{\mathbf{k}}} a_{kj}^{\mathbf{k}} \quad \forall j = k, \dots, n+1 \right) \quad \forall i = k+1, \dots, n \right) \quad \forall k = 1, \dots, n-1$$

Il codice ottimizzato:

$$\begin{aligned} &\forall k = 1, \dots, n-1 \\ &\quad \forall i = k+1, \dots, n \\ &\quad \quad m := \frac{a_{ik}^{\mathbf{k}}}{a_{kk}^{\mathbf{k}}} \\ &\quad \quad \forall j = k, \dots, n+1 \\ &\quad \quad \quad a_{ij}^{\mathbf{k}+1} := a_{ij}^{\mathbf{k}} - m a_{kj}^{\mathbf{k}} \end{aligned}$$

(1) Calcoliamo il costo

$$\begin{aligned} [a_{ij}^{\mathbf{k}+1}] &= [a_{ij}^{\mathbf{k}} - m a_{kj}^{\mathbf{k}}] = [m a_{kj}^{\mathbf{k}}] = 1 \\ I_i &:= [a_{ij}^{\mathbf{k}+1} \quad \forall j = k, \dots, n+1] = \sum_{j=k}^{n+1} [a_{ij}^{\mathbf{k}+1}] = n+1 - k + 1 = n - k + 2 \\ I'_i &:= \left[m := \frac{a_{ik}^{\mathbf{k}}}{a_{kk}^{\mathbf{k}}} \right] + K_i = 1 + K_i = n - k + 3 \\ K_k &= \sum_{i=k+1}^n I'_i = \sum_{i=k+1}^n n - k + 3 = (n - k + 3)(n - k) \\ K &= \sum_{k=1}^{n-1} K_k = \sum_{k=1}^{n-1} (n - k + 3)(n - k) \underbrace{=}_{l=n-k} = \sum_{l=1}^{n-1} (l+3)l = 3 \sum_{l=1}^{n-1} l + \sum_{l=1}^{n-1} l^2 = \\ &\underbrace{=}_{[2.1, \text{pg.2}], [2.2, \text{pg.2}]} 3 \frac{(n-1)n}{2} + \frac{2(n-1)^3 + 3(n-1)^2 + n-1}{6} \end{aligned}$$

□

Proposition 5.3. Ponendoci nel contesto della prop [5.2,pg.15], si ha

$$\exists k : a_{kk}^{\mathbf{k}} = 0 \Leftrightarrow \exists k : \det D_k = 0$$

dove D_k e' il minore principale di A di ordine k .

Equivalentemente:

$$\forall k = 1, \dots, n \quad a_{kk}^{\mathbf{k}} \neq 0 \Leftrightarrow \forall k = 1, \dots, n \quad \det D_k \neq 0$$

Esempi in cui si verificano queste condizioni sono:

A simmetrica, definita positiva, ovvero

$$xA^t x > 0 \quad \forall x \in \mathbb{R}^{1,n}, \quad A = {}^t A$$

A simmetrica, definita negativa, ovvero

$$xA^t x < 0 \quad \forall x \in \mathbb{R}^{1,n}, \quad A = {}^t A$$

A a predominanza diagonale stretta, ovvero

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

Proof:

(1) Dim. $\Rightarrow$

Supponiamo che $a_{hh}^{\mathbf{h}} = 0$ sia il primo di quelli nulli, ovvero, supponiamo che

$$a_{ii}^{\mathbf{i}} \neq 0 \quad \forall i = 1, \dots, h-1$$

Possiamo allora applicare l'algoritmo [5.2,pg.15] fino al passo h :

$$\left(A_i^{\mathbf{k}+1} := A_i^{\mathbf{k}} - \frac{a_{ik}^{\mathbf{k}}}{a_{kk}^{\mathbf{k}}} A_k^{\mathbf{k}} \quad \forall i = k+1, \dots, n \right) \quad \forall k = 1, \dots, h-1$$

Poiche' ogni nuova matrice e' ottenuta aggiungendo alle righe della precedente dei multipli di altre sue righe, per le proprieta' del determinante, si ha

$$\det D_k^{\mathbf{h}} = \det D_k^{\mathbf{1}} = \det D_k \quad \forall k = 1, \dots, n$$

dove con $D_k^{\mathbf{h}}$ intendiamo il minore principale di ordine k della matrice $A^{\mathbf{h}}$.

Ma $D_h^{\mathbf{h}}$ e' una matrice triangolare, e quindi il suo determinante e'

$$D_h^{\mathbf{h}} = a_{11}^{\mathbf{h}} a_{22}^{\mathbf{h}} \dots a_{hh}^{\mathbf{h}} \underbrace{=}_{{a_{hh}^{\mathbf{h}}=0}} 0$$

(1) Dim. $\Leftarrow$

Sia $\det D_h = 0$, e

$$\det D_i \neq 0 \quad \forall i = 1, \dots, h-1$$

Se $\exists i < h : a_{ii}^{\mathbf{i}} = 0$, allora non abbiamo nulla da dimostrare. Supponiamo allora che

$$a_{ii}^{\mathbf{i}} \neq 0 \quad \forall i = 1, \dots, h-1 \quad (1)$$

Procedendo come nel passo precedente, si ha:

$$a_{11}^{\mathbf{h}} a_{22}^{\mathbf{h}} \dots a_{hh}^{\mathbf{h}} = \det D_h^{\mathbf{h}} = \det D_h = 0 \underbrace{\Rightarrow}_{(1)} a_{hh}^{\mathbf{h}} = 0$$

(1) Esempi

Se A simmetrica e' definita positiva, per la caratterizzazione [GII,2.56,pg.56], si ha

$$D_1, D_2, \dots, D_n > 0$$

se invece e' definita negativa si ha

$$D_1 < 0, D_2 > 0, D_3 < 0, \dots$$

In entrambi i casi, nessun minore e' nullo.

□

Proposition 5.4. Sia $A \in \mathbb{R}^{n \times n}$, $|A| \neq 0$, $\mu(A) = \|A^{-1}\| \|A\|$. Considerando il sistema lineare

$$Ax = b$$

e il sistema perturbato da un errore

$$A(x + \Delta x) = b + \Delta b$$

si ha

$$\frac{\|\Delta x\|}{\|x\|} \leq \mu(A) \frac{\|\Delta b\|}{\|b\|}$$

Ovvero, l'errore relativo sul risultato e' limitato dall'errore relativo sui termini noti, per un fattore di $\mu(A)$. $\mu(A)$ viene chiamato numero di condizionamento di A . Usando la norma 2, si ha

$$\mu(A) = \frac{\sqrt{\lambda_M({}^tAA)}}{\sqrt{\lambda_m({}^tAA)}} \\ A = {}^tA \Rightarrow \mu(A) = \frac{|\lambda_M|}{|\lambda_m|}$$

Proof:

(1) Dim 1.

Osserviamo che

$$b = Ax \Rightarrow \|b\| = \|Ax\| \underbrace{\leq}_{[4.8,pg.11]} \|A\| \|x\| \Rightarrow \frac{\|b\|}{\|x\|} \leq \|A\| \quad (1)$$

$$A(x + \Delta x) = b + \Delta b \Leftrightarrow Ax + A\Delta x = b + \Delta b \Leftrightarrow A\Delta x = \Delta b \Leftrightarrow \Delta x = A^{-1}\Delta b$$

$$\|\Delta x\| = \|A^{-1}\Delta b\| \underbrace{\leq}_{[4.8,pg.11]} \|A^{-1}\| \|\Delta b\| \Leftrightarrow \frac{\|\Delta x\|}{\|x\|} \leq \|A^{-1}\| \frac{\|\Delta b\|}{\|b\|} = \|A^{-1}\| \frac{\|\Delta b\|}{\|b\|} \frac{\|b\|}{\|x\|} \leq$$

$$\underbrace{\leq}_{(1)} \|A^{-1}\| \|A\| \frac{\|\Delta b\|}{\|b\|} = \mu(A) \frac{\|\Delta b\|}{\|b\|}$$

(2) Dim 2.

Poiche' $\mu_2(A) = \|A\|_2 \|A^{-1}\|_2$, basta usare le prop [4.10,pg.12], [4.11,pg.13].

□

5.1 Fattorizzazione LU

Theorem 5.5. Data $A \in \mathbb{R}^{n \times n}$, si ha

$$D_k \neq 0 \quad \forall k = 1, \dots, n \Rightarrow A = LU$$

dove L, U sono le seguenti matrici triangolari :

$$L = \begin{bmatrix} 1 & & & & \\ m_{21} & 1 & & & \\ m_{31} & m_{32} & 1 & & \\ \dots & & & \ddots & \\ m_{n1} & m_{n2} & \dots & m_{n,n-1} & 1 \end{bmatrix}, \quad U = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22}^2 & a_{23} & \dots & a_{2n}^2 \\ 0 & 0 & a_{33}^3 & \dots & a_{3n}^3 \\ \dots & & & \ddots & \\ 0 & 0 & 0 & \dots & a_{nn}^n \end{bmatrix}$$

Dove m_{ij}, a_{ij}^k sono ottenuti tramite l'algoritmo [5.2,pg.15].

Proof: Per la prop [5.3,pg.18], la condizione $D_k \neq 0 \quad \forall k = 1, \dots, n$ assicura che e' possibile utilizzare l'algoritmo [5.2,pg.15], ovvero

$$a_{kk}^k \neq 0 \quad \forall k = 1, \dots, n$$

Sia

$$\underline{m}_i = \begin{pmatrix} 0 & 0 & \dots & \underbrace{0}_{\text{posto } i\text{-esimo}} & m_{i+1,i} & m_{i+2,i} & \dots & m_{n,i} \end{pmatrix}$$

$$\underline{m}_{ij} = (\underline{m}_i)_j = \begin{cases} 0 & j \leq i \\ m_{ji} & j > i \end{cases} \quad (1)$$

$$M_i = I - {}^t \underline{m}_i e_i$$

con $\{e_i\}$ base canonica di $\mathbb{R}^n$.

(1) Si ha

$$(M_h)_{ij} = \begin{cases} 1 & i = j \\ -m_{ji} & i > h, j = h \\ 0 & \text{altrimenti} \end{cases}$$

Ovvero, M_h e' la matrice triangolare che ha tutti 1 nella diagonale principale, come colonna h ha $-{}^t \underline{m}_i$, tutti gli altri elementi sono 0:

$$M_h = \begin{bmatrix} 1 & & & & & & & \\ 0 & 1 & & & & & & \\ \vdots & & \ddots & & & & & \\ 0 & 0 & \dots & 0 & 1 & & & \\ 0 & 0 & \dots & 0 & -m_{k+1,k} & & & \\ 0 & 0 & \dots & 0 & -m_{k+2,k} & & & \\ \vdots & & & & & \ddots & & \\ 0 & 0 & \dots & 0 & \underbrace{-m_{n,k}}_{\text{colonna } k} & & 1 \end{bmatrix}$$

Osserviamo che poiche' ${}^t \underline{m}_h \in \mathbb{R}^{n \times 1}$, $e_h \in \mathbb{R}^{1 \times n}$, si ha ${}^t \underline{m}_h e_h \in \mathbb{R}^{n \times n}$. Ed e'

$$({}^t \underline{m}_h e_h)_{ij} = {}^t \underline{m}_{hi} e_{hj} = \underbrace{{}^t \underline{m}_{hi} \delta_{hj}}_{(1)} = \begin{cases} 0 & i \leq h \\ m_{ih} \delta_{hj} & i > h \end{cases} = \begin{cases} m_{ih} & i > h, j = h \\ 0 & \text{altrimenti} \end{cases}$$

$$(M_h)_{ij} = (I - {}^t \underline{m}_h e_h)_{ij} = \delta_{ij} - \begin{cases} m_{ih} & i > h, j = h \\ 0 & \text{altrimenti} \end{cases} = \begin{cases} 1 & i = j \\ -m_{ih} & i > h, j = h \\ 0 & \text{altrimenti} \end{cases}$$

(2) Si ha $M_{h-1} \dots M_1 A = D_n^h$

Dove D_k^h e' il minore principale di ordine k di A^h .

Consideriamo $M_1 A$

$$(M_1 A)_{ij} = \sum_{r=1}^n (M_1)_{ir} A_{rj} \underbrace{=}_{\text{passo 1}} \sum_{r=1}^n \begin{cases} A_{rj} & i = r \\ -m_{ri} A_{rj} & i > 1, r = 1 \\ 0 & \text{altrimenti} \end{cases}$$

$$M_1 A = \begin{bmatrix} 1 & & & \\ -m_{21} & 1 & & \\ -m_{31} & 0 & 1 & \\ \vdots & & & \\ -m_{n1} & & & 1 \end{bmatrix} \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \\ a_{n1} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} - m_{21}a_{11} & \dots & a_{2n} - m_{21}a_{1n} \\ a_{31} - m_{31}a_{11} & \dots & a_{3n} - m_{31}a_{1n} \\ \vdots & & \\ a_{n1} - m_{n1}a_{11} & \dots & a_{nn} - m_{n1}a_{1n} \end{bmatrix}$$

E ricordando che m_{ij} e' il moltiplicatore di posto i, j si ha

$$M_1 A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} - m_{21}a_{12} & \dots & a_{2n} - m_{21}a_{1n} \\ \vdots & & & \\ 0 & a_{n2} - m_{n1}a_{12} & \dots & a_{nn} - m_{n1}a_{1n} \end{bmatrix} = \begin{bmatrix} a_{11}^2 & a_{12}^2 & a_{13}^2 & \dots & a_{1n}^2 \\ 0 & a_{22}^2 & a_{23}^2 & \dots & a_{2n}^2 \\ \vdots & & & & \\ 0 & a_{n2}^2 & a_{n3}^2 & \dots & a_{nn}^2 \end{bmatrix} = D_n^2$$

Così procedendo si arriva a

$$M_{h-1} \dots M_1 A = D_n^h$$

Sia $N = M_{n-1} \dots M_1$, Sia $U = NA$. Si ha

$$U \underbrace{=}_{\text{passo 2}} D_n^n$$

D_n^n e' il minore di ordine n dell'ultima matrice data dall'algoritmo [5.2,pg.15]. Quindi, U coincide con l' U della tesi.

(3) $\exists N^{-1}$

Abbiamo già visto che M_i e' una matrice triangolare. Poiché gli elementi della sua diagonale sono tutti 1, si ha

$$|M_i| = 1 \Rightarrow \exists M_i^{-1} \quad \forall i = 1, \dots, n$$

$$N^{-1} = M_1^{-1} M_2^{-1} \dots M_{n-1}^{-1}$$

Sia $L = N^{-1}$, si ha

$$U = NA \Rightarrow A = N^{-1}U = LU$$

(4) Resta da controllare che L sia proprio quello della tesi

$$\langle 4.1 \rangle \quad M^{-1}_i = I + {}^t \underline{m}_i e_i$$

Infatti,

$$M_i(I + {}^t \underline{m}_i e_i) = (I - {}^t \underline{m}_i e_i)(I + {}^t \underline{m}_i e_i) = I - {}^t \underline{m}_i e_i {}^t \underline{m}_i e_i =$$

$$e_i {}^t \underline{m}_i = e_i \cdot \underline{m}_i = \underline{m}_{ii} \underbrace{=}_{(1)} 0$$

$$= I$$

$$(I + {}^t \underline{m}_i e_i) M_i = I$$

$$\langle 4.2 \rangle \quad L = I + \sum_{i=1}^{n-1} {}^t \underline{m}_i e_i$$

$$\begin{aligned}
L &= N^{-1} = M_1^{-1} M_2^{-1} \dots M_{n-1}^{-1} = (I + {}^t \underline{m}_1 e_1)(I + {}^t \underline{m}_2 e_2) \dots (I + {}^t \underline{m}_{n-1} e_{n-1}) \\
(I + {}^t \underline{m}_1 e_1)(I + {}^t \underline{m}_2 e_2) &= I + {}^t \underline{m}_1 e_1 + {}^t \underline{m}_2 e_2 + {}^t \underline{m}_1 e_1 {}^t \underline{m}_2 e_2 = \\
&= e_1 {}^t \underline{m}_2 = e_1 \cdot \underline{m}_2 = \underbrace{\underline{m}_{21}}_{(1)} = 0 \\
&= I + {}^t \underline{m}_1 e_1 + {}^t \underline{m}_2 e_2 \\
(I + {}^t \underline{m}_1 e_1 + \dots + {}^t \underline{m}_{n-2} e_{n-2})(I + {}^t \underline{m}_{n-1} e_{n-1}) &= I + \sum_{i=1}^{n-1} {}^t \underline{m}_i e_i + \sum_{i=1}^{n-2} {}^t \underline{m}_i e_i {}^t \underline{m}_{n-1} e_{n-1} = \\
e_i {}^t \underline{m}_{n-1} &= e_i \cdot \underline{m}_{n-1} = \underbrace{\underline{m}_{n-1,i}}_{n-1 \geq i, (1)} = 0 \\
&= I + \sum_{i=1}^{n-1} {}^t \underline{m}_i e_i
\end{aligned}$$

(4.3) Q.E.D.

E considerando che le matrici ${}^t \underline{m}_i e_i$ sono del tipo:

$${}^t \underline{m}_i e_i = \begin{bmatrix} 0 & & & & & \\ 0 & 0 & & & & \\ \vdots & & & & & \\ 0 & 0 & \dots & 0 & 0 & \\ 0 & 0 & \dots & 0 & m_{i+1,i} & \\ 0 & 0 & \dots & 0 & m_{i+2,i} & \\ \vdots & & & & & \\ 0 & 0 & \dots & 0 & \underbrace{m_{ni}}_{\text{colonna i}} & 0 \end{bmatrix}$$

si ha che $L = I + \sum_{i=1}^{n-1} {}^t \underline{m}_i e_i$ coincide con l' L della tesi.

□

Proposition 5.6. *Poniamoci nel contesto della prop [5.2,pg.15], senza supporre che $a_{kk}^{\mathbf{k}} \neq 0 \ \forall k = 1, \dots, n-1$.*

Sia $s := (1, 2, \dots, n)$, inizializzato all'inizio dell'algoritmo.

Definiamo le seguenti procedure:

Sia scambio_righe i, i' :

$$\begin{cases} T := a_{ij} \\ a_{ij} := a_{i'j} \\ a_{i'j} := T \end{cases} \quad \forall j := 1, \dots, n$$

Sia scambio_colonne j, j' :

$$\begin{cases} T := a_{ij} \\ a_{ij} := a_{ij'} \\ a_{ij'} := T \end{cases} \quad \forall j := 1, \dots, n$$

$$\begin{cases} t := s_j \\ s_j := s_{j'} \\ s_{j'} := t \end{cases}$$

Sia pivot_riga k :

$$\mathbf{Sia} \ \bar{i} \ \text{tale che} \ |a_{\bar{i}k}| := \max_{i:=k, \dots, n} |a_{ik}|$$

scambio_righe $\bar{i}, k$

Sia pivot_colonna k :

$$\mathbf{Sia} \ \bar{j} \ \text{tale che} \ |a_{k\bar{j}}| := \max_{j:=k, \dots, n} |a_{kj}|$$

scambio_colonne $\bar{j}, k$

Sia pivot_rigcol k :

pivot_riga k

pivot_colonna k

Sia pivot_scalato k :

$$d_i := \frac{|a_{ik}|}{\max_{j=1, \dots, n} |a_{ij}|} \quad \forall i = k, \dots, n$$

$$\mathbf{Sia} \ \bar{i} \ \text{tale che} \ d_{\bar{i}} = \max_{i=k, \dots, n} d_i$$

scambio_riga $\bar{i}, k$

Sia ripristina_colonne :

$$\begin{cases} \mathbf{Se} \ i \neq s_i : \\ \text{scambio_colonna } i, s_i \end{cases} \quad \forall i = 1, \dots, n$$

L'algoritmo di eliminazione di Gauss generale, e' il seguente:

```

s := (1, 2, ..., n)
∀k = 1, ..., n - 1
  pivot k
  ∀i = k + 1, ..., n
    m :=  $\frac{a_{ik}^{\mathbf{k}}}{a_{kk}^{\mathbf{k}}}$ 
    ∀j = k + 1, ..., n + 1
       $a_{ij}^{\mathbf{k}+1} := a_{ij}^{\mathbf{k}} - ma_{kj}^{\mathbf{k}}$ 
  Se pivot = pivot_colonna ∨ pivot = pivot_rigcol :
    ripristina_colonne

```

Nota: pivot puo' essere una qualsiasi delle procedure di pivot sopra definite.
 Infine, l'algoritmo che permette di usare una sola matrice in memoria, e nessuna variabile ausiliaria per i moltiplicatori, durante tutto il corso dell'esecuzione e':

```

s := (1, 2, ..., n)
∀k = 1, ..., n - 1
  pivot k
  ∀i = k + 1, ..., n
     $a_{ik} := \frac{a_{ik}}{a_{kk}}$ 
    ∀j = k + 1, ..., n + 1
       $a_{ij} := a_{ij} - a_{ik}a_{kj}$ 
  Se pivot = pivot_colonna ∨ pivot = pivot_rigcol :
    ripristina_colonne

```

Proof: Per costruzione. la scelta dei vari pivot $(\bar{i}, \bar{j})$, permette di scambiare gli $a_{kk}^{\mathbf{k}} = 0$ con altri coefficienti non nulli. Questo avviene tramite lo scambio di righe e/o colonne. Ovviamente, lo scambio di righe o colonne effettuato da pivot non altera il sistema. Queste tecniche di pivoting, permettono anche di evitare di incorrere nella situazione in cui il moltiplicatore m diventa molto grande. In questo caso nella differenza

$$a_{ij} := a_{ij} - ma_{kj} \quad (1)$$

si avrebbe una perdita sostanziale di cifre significative.

I pivot pivot_rigcol, pivot_scalato permettono anche di evitare la situazione in cui nella (1), a_{ij} e' molto piu' piccolo rispetto a a_{kj} .

Infine, l'algoritmo ottimizzato per la memoria usa la cella della matrice che sara' sovrascritta da uno zero per memorizzare il moltiplicatore, e inoltre, sfrutta il fatto che $a_{ij}^{\mathbf{k}} = a_{ij}^{\mathbf{k}+1} = \dots = a_{ij}^{\mathbf{n}}$. □

Theorem 5.7. Data $A \in \mathbb{R}^{n \times n}$, $|A| \neq 0$, si ha

$$\exists P \text{ matrice di permutazione: } PA = LU$$

dove L, U sono le matrici triangolari descritte nel thm [5.5,pg.19], che si ottengono pero' usando l'algoritmo [5.6,pg.22] con un pivot che non scambia colonne.

Nota: e' essenziale usare la versione dell'algoritmo che salva i moltiplicatori nella stessa matrice A, infatti, anche le righe di L devono essere scambiate durante l'esecuzione.

Una matrice P, si dice matrice di permutazione se e' ottenuta dalla matrice identica I, mediante scambi di righe (o, equivalentemente, di colonne).

Proof: Ricordiamo che PA e' quella matrice ottenuta da A effettuando lo stesso scambio di righe che sono state utilizzate per ottenere P da I (vedi [AL,1.4.1,pg.1]).

Rileggendo l'algoritmo [5.6,pg.22], possiamo esprimere lo scambio di righe (o di colonne) di A mediante una opportuna matrice di permutazione P_i . I passi compiuti dall'algoritmo si possono allora esprimere cosi':

$$\begin{aligned} M_1 P_1 A x &= M_1 P_1 b \\ M_2 P_2 M_1 P_1 A x &= M_2 P_2 M_1 P_1 b \\ &\dots \\ M_{n-1} P_{n-1} \dots M_2 P_2 M_1 P_1 A x &= M_{n-1} P_{n-1} M_2 P_2 M_1 P_1 b \end{aligned}$$

Consideriamo la matrice $M_{n-1} P_{n-1} \dots M_2 P_2 M_1 P_1 A$ e poniamo $P = P_{n-1} P_{n-2} \dots P_1$. Esaminiamo solo il caso $n = 4$.

$$\begin{aligned} U &= (M_3 P_3)(M_2 P_2)(M_1 P_1) A \underbrace{=}_{P_i P_i = I} (M_3 P_3)(M_2 P_3 P_3 P_2)(M_1 P_2 P_3 P_3 P_2 P_1) A = \\ &= M_3 (P_3 M_2 P_3)(P_3 P_2 M_1 P_2 P_3)(P_3 P_2 P_1) A = L^{-1} P A \Leftrightarrow P A = L U \end{aligned}$$

□

5.1.1 Applicazioni

Proposition 5.8. Data la matrice $A \in \mathbb{R}^{n \times n}$, $|A| \neq 0$, e la matrice U ottenuta dalla fattorizzazione [5.7,pg.24], si ha

$$\det A = (-1)^s \prod_{i=1}^n u_{ii}$$

dove s e' il numero di scambi di righe effettuati durante la fattorizzazione [5.7,pg.24], o equivalentemente $(-1)^s = \det P$

Il costo di questo calcolo e' quindi

$$\text{costo di [5.7,pg.24]} + n - 1 \approx \frac{n^3}{3} + O(n^2)$$

Proof: Si ha

$$\det PA = \det P \det A = \det P \det A$$

$$\det PA = \det LU = \det L \det U = \det U \underbrace{=}_{U \text{ e' triangolare}} \prod_{i=1}^n u_{ii}$$

$$\det P \det A = \prod_{i=1}^n u_{ii}$$

Ricordando che uno scambio di riga cambia di segno il determinante, si ha

$$\det P = (-1)^s \det I = (-1)^s$$

$$\det A = (-1)^s \prod_{i=1}^n u_{ii}$$

dove s e' il numero di righe che sono state scambiate per ottenere P da I . Ma dato che $P = P_{n-1} \dots P_1$, s si puo' conoscere direttamente considerando il numero di scambi di righe effettuati durante l'uso dell'algoritmo [5.7,pg.24]. $\square$

Proposition 5.9. *Sia $A \in \mathbb{R}^{n \times n}$, $|A| \neq 0$. Siano dati m sistemi lineari del tipo:*

$$Ax_k = b_k \quad k = 1, \dots, m$$

Per risolverli si puo' procedere cosi':

Si fattorizza A in $PA = LU$

$$\forall k = 1, \dots, m$$

$$y := Ux_k$$

$$\bar{b}_k = Pb_k$$

Si risolve con [5.1,pg.13] il sistema triangolare $Ly = \bar{b}_k$

Si risolve con [5.1,pg.13] il sistema triangolare $Ux_k = y$

Il costo di questo algoritmo e'

$$\frac{n^3}{3} + mn^2 + O(n)$$

Proof: Fissiamo k

$$Ax_k = b_k$$

$$PAx_k = Pb_k$$

$$LUx_k = Pb_k = \bar{b}_k$$

$$\text{Sia } y = Ux_k$$

$$Ly = \bar{b}_k$$

Il costo dell'algoritmo [5.1,pg.13] usato per risolvere il sistema triangolare e' $\frac{n^2}{2} + O(n)$. Poiche' stiamo risolvendo due sistemi triangolari, abbiamo $n^2 + O(n)$. $\square$

Proposition 5.10. *Sia $A \in \mathbb{R}^{n \times n}$, $|A| \neq 0$. Sia $C = A^{-1}$.*

Per calcolare C si puo' procedere cosi':

$$\forall i = 1, \dots, n$$

Si risolve con [5.9,pg.26] il sistema $AC^i = e_i$

dove C^i e' la colonna i -esima di C .

Il costo e'

$$\frac{4}{3}n^3 + O(n)$$

Proof: $AC = AA^{-1} = I$, e poiche' $AC^i = I^i$, si deve avere

$$AC^i = e_i \quad \forall i = 1, \dots, n$$

Applicando la [5.9,pg.26], con $m = n$, si ricava il costo dell'algoritmo. $\square$

5.2 Fattorizzazione di Cholesky

Theorem 5.11. Sia $A \in \mathbb{R}^{n \times n}$ una matrice simmetrica e definita positiva, allora

$$\exists R \in \mathbb{R}^{n \times n} \text{ triangolare inferiore: } A = R^t R$$

L'algoritmo per ricavare R e':

$$\begin{aligned} \forall j = 1, \dots, n \\ r_{jj} &:= \sqrt{a_{jj} - \sum_{k=1}^{j-1} r_{jk}^2} \\ \forall i = j+1, \dots, n \\ r_{ij} &:= \frac{a_{ij} - \sum_{k=1}^{j-1} r_{ik} r_{jk}}{r_{jj}} \end{aligned}$$

Proof: La condizione $A = R^t R$ si traduce nel sistema

$$\begin{aligned} a_{ij} &= \sum_{k=1}^n R_{ik} ({}^t R)_{kj} \quad \forall j = 1, \dots, n, \quad \forall i = 1, \dots, n \Leftrightarrow \\ &\Leftrightarrow a_{ij} = \sum_{k=1}^n R_{ik} R_{jk} \quad \forall j = 1, \dots, n, \quad \forall i = 1, \dots, n \end{aligned}$$

dove r_{hk} sono gli elementi di R e le incognite del sistema ⁵.

Tenendo conto del fatto che R e' una matrice triangolare inferiore, possiamo scrivere

$$a_{ij} = \sum_{k=1}^{\min\{i,j\}} R_{ik} ({}^t R)_{kj} \quad \forall j = 1, \dots, n, \quad \forall i = 1, \dots, n$$

Poiche' A e' simmetrica, $a_{ij} = a_{ji}$, quindi le uniche equazioni che ci interessano sono:

$$a_{ij} = \sum_{k=1}^{\min\{i,j\}} r_{ik} r_{jk} \quad \forall j = 1, \dots, i, \quad \forall i = 1, \dots, n$$

Adesso basta scrivere le equazioni nel seguente ordine:

$$\begin{cases} 1. & a_{11} = r_{11}^2 \\ 2. & a_{i1} = r_{i1} r_{11} \quad \forall i = 2, \dots, n \\ 3. & a_{22} = r_{21}^2 + r_{22}^2 \\ 4. & a_{i2} = r_{i1} r_{21} + r_{i2} r_{22} \quad \forall i = 3, \dots, n \\ \vdots & \end{cases}$$

e ricavare le equazioni dall'alto verso il basso. Per ricordare l'ordine delle equazioni si puo' usare questo disegno:

$$\begin{bmatrix} 1. & & \\ 2. & 3. & \\ 2. & 4. & 5. \\ 2. & 4. & 6. \end{bmatrix}$$

⁵Sfruttando le ipotesi fatte su A , si puo' dimostrare per via teorica che tale sistema ammette soluzione

Le incognite sono:

$$\begin{aligned} r_{11} &= \sqrt{a_{11}} \\ r_{i1} &= \frac{a_{i1}}{r_{11}} \quad \forall i = 2, \dots, n \\ r_{22} &= \sqrt{a_{22} - r_{21}^2} \\ r_{i2} &= \frac{a_{i2} - r_{i1}r_{21}}{r_{22}} \quad \forall i = 3, \dots, n \\ &\dots \end{aligned}$$

Le radici quadrate sono lecite perche' A e' definita positiva.

In generale,

$$\begin{aligned} a_{jj} &= \sum_{k=1}^j r_{jk}^2 = r_{jj}^2 + \sum_{k=1}^{j-1} r_{jk}^2 \Leftrightarrow r_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} r_{jk}^2} \\ a_{ij} &= \sum_{k=1}^j r_{ik}r_{jk} = r_{ij}r_{jj} + \sum_{k=1}^{j-1} r_{ik}r_{jk} \Leftrightarrow r_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} r_{ik}r_{jk}}{r_{jj}} \quad \forall i = j+1, \dots, n \end{aligned}$$

L'algoritmo completo e' quindi

$$\begin{aligned} &\forall j = 1, \dots, n \\ &r_{jj} := \sqrt{a_{jj} - \sum_{k=1}^{j-1} r_{jk}^2} \\ &\forall i = j+1, \dots, n \\ &r_{ij} := \frac{a_{ij} - \sum_{k=1}^{j-1} r_{ik}r_{jk}}{r_{jj}} \end{aligned}$$

□

5.3 Autovalori

Theorem 5.12. Data $A \in \mathbb{R}^{m \times n}$, con $m \geq n$,

$\exists U \in \mathbb{R}^{m \times m}, \Sigma \in \mathbb{R}^{m \times n}, V \in \mathbb{R}^{n \times n}$ ortogonali, tali che $A = U\Sigma^tV$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \sigma_3 & & \\ \dots & & & & \\ & & & & \sigma_n \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$$

I σ_i si chiamano valori singolari.

Proposition 5.13. *Nel contesto del thm [5.12,pg.28], se r e' il rango di A , allora*

$$A = \sum_{i=1}^r \sigma_i U^i V^i$$

dove U^i e' la colonna i -esima di U , V^i la colonna i -esima di V . Inoltre,

$$\sigma_1 = \|A\|_2$$

Proof: Per il thm [5.12,pg.28]

$$A = U \Sigma^t V = \sum_{i=1}^n \sigma_i U_i^t V^i$$

Si ha

$$\sigma_h \neq 0 \quad \underbrace{\Rightarrow}_{\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0} \quad \sigma_1, \dots, \sigma_h \neq 0$$

□

Theorem 5.14. *di Gershgorin-Hadamard. Sia $A \in \mathbb{C}^{n \times n}$, e siano*

$$D_i = \{z \in \mathbb{C} \mid |z - a_{ii}| \leq R_i\} \quad [e' \text{ un cerchio in } \mathbb{C} \text{ di centro } a_{ii}, \text{ raggio } R_i]$$

$$R_i = \sum_{j \neq i} |a_{ij}|$$

$$D = \bigcup_{i=1}^n D_i$$

Si ha:

$$\lambda \text{ autovalore di } A \Rightarrow \lambda \in D$$

Proof: Sia $u \in \mathbb{C}^{n \times 1}$ l'autovettore di A , associato a λ . Si ha

$$u_k \text{ autovettore} \Rightarrow u_k \neq 0$$

$$Au = \lambda u \Leftrightarrow \sum_j a_{ij} u_j = \lambda u_i \Leftrightarrow a_{ii} u_i + \sum_{j \neq i} a_{ij} u_j = \lambda u_i \Leftrightarrow \sum_{j \neq i} a_{ij} u_j = (\lambda - a_{ii}) u_i$$

$$|\sum_{j \neq i} a_{ij} u_j| = |(\lambda - a_{ii}) u_i| = |\lambda - a_{ii}| |u_i|$$

$$|\lambda - a_{ii}| |u_i| = |\sum_{j \neq i} a_{ij} u_j| \leq \sum_{j \neq i} |a_{ij}| |u_j| \quad (1)$$

Sia $k : |u_k| = \max \{|u_i|\}_i$. Ponendo $i = k$ nella (1), si ha

$$|\lambda - a_{kk}| |u_k| \leq \sum_{j \neq k} |a_{kj}| |u_j| \quad \underbrace{\leq}_{|u_k| = \max \{|u_i|\}} \sum_{j \neq k} |a_{kj}| |u_k| = |u_k| \sum_{j \neq k} |a_{kj}|$$

$$|\lambda - a_{kk}| |u_k| \leq |u_k| \sum_{j \neq k} |a_{kj}| \quad \underbrace{\Rightarrow}_{u_k \neq 0} |\lambda - a_{kk}| \leq \sum_{j \neq k} |a_{kj}| = R_k \Rightarrow \lambda \in D_k \subseteq D$$

□

Proposition 5.15. *Metodo delle potenze per il calcolo dell'autovalore massimo. Siano $\lambda_1, \dots, \lambda_n$ tutti gli autovalori di $A \in \mathbb{C}^{n \times n}$, matrice diagonalizzabile. Sia*

$$\begin{aligned} x_0 &\in \mathbb{C}^{n,1}, \quad x_0 \neq 0 \\ \begin{cases} y_k = \frac{x_k}{\|x_k\|} \\ x_{k+1} = Ay_k \end{cases} &\quad \forall k = 0, 1, 2, \dots \\ \sigma_k &= \frac{{}^t x_k A x_k}{{}^t x_k x_k} \end{aligned}$$

$\sigma(x) = \frac{{}^t x A x}{{}^t x x}$ e' chiamato rapporto di Rayleigh. Supponiamo che

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$$

Allora si ha:

$$\begin{aligned} \sigma_k &= \frac{{}^t y_k x_{k+1}}{{}^t y_k y_k} \\ \lim_{k \rightarrow \infty} \sigma_k &= \lambda_1 \end{aligned}$$

Una successione piu' semplice, ma meno stabile dal punto di vista numerico e'

$$z_0 = x_0, \quad z_{k+1} = Az_k, \quad \sigma'_k = \frac{{}^t z_k A z_k}{{}^t z_k z_k}$$

Supponendo che

$$|\lambda_1| > |\lambda_2| > |\lambda_3| > \dots > |\lambda_h| > |\lambda_{h+1}| \geq |\lambda_{h+2}| \geq \dots \geq |\lambda_n|$$

per ricavare λ_h poniamo

$$x'_0 = x_0 - \sum_{i=1}^{h-1} \alpha_i u_i$$

,dove α_i, u_i sono descritti nella dimostrazione, e riappliciamo lo stesso algoritmo. Il nuovo σ_k convergera' a λ_h .

Infine, scelto un errore $\varepsilon > 0$, possiamo arrestare la successione quando

$$\frac{\|(A - \sigma_k I)y_k\|}{\sigma_k \|y_k\|} < \varepsilon$$

oppure quando

$$|\sigma_{k+1} - \sigma_k| < \varepsilon$$

Proof: Sia $z_0 = x_0$, $z_{k+1} = Az_k$, si ha
 A diagonalizzabile $\Rightarrow \exists B = \{u_1, \dots, u_n\}$ base di $\mathbb{C}^{n \times n}$: $Au_i = \lambda_i u_i$ (1)

$$\begin{aligned} B \text{ base} &\Rightarrow x_0 = \sum_{i=1}^n \alpha_i u_i \\ z_1 &= Az_0 = \sum_{i=1}^n \alpha_i Au_i \underbrace{=}_{(1)} \sum_{i=1}^n \alpha_i \lambda_i u_i \\ z_2 &= Az_1 = \sum_{i=1}^n \alpha_i \lambda_i Au_i \underbrace{=}_{(1)} \sum_{i=1}^n \alpha_i \lambda_i^2 u_i \\ &\dots \\ z_k &= Az_{k-1} = \sum_{i=1}^n \alpha_i \lambda_i^k u_i \quad (1.1) \end{aligned}$$

Con $[v]_i$ indichiamo la componente i -esima del vettore v secondo la base canonica. Se $\alpha_1 u_{1l} \neq 0$, si ha

$$\begin{aligned} \frac{[z_{k+1}]_l}{[z_k]_l} &= \frac{\sum_{i=1}^n \alpha_i \lambda_i^{k+1} u_{il}}{\sum_{i=1}^n \alpha_i \lambda_i^k u_{il}} = \frac{\lambda_1^{k+1}}{\lambda_1^k} \frac{\alpha_1 u_{1l} + \sum_{i=2}^n \alpha_i \frac{\lambda_i^{k+1}}{\lambda_1^{k+1}} u_{il}}{\alpha_1 u_{1l} + \sum_{i=2}^n \alpha_i \frac{\lambda_i^k}{\lambda_1^k} u_{il}} = \\ &= \lambda_1 \frac{\alpha_1 u_{1l} + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k+1} u_{il}}{\alpha_1 u_{1l} + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^k u_{il}} \quad (2) \end{aligned}$$

Poiche' $\lambda_1 > \lambda_i$ si ha che

$$\lim_{k \rightarrow \infty} \left(\frac{\lambda_i}{\lambda_1}\right)^k = 0$$

e quindi dalla (2)

$$\lim_{k \rightarrow \infty} \frac{[z_{k+1}]_l}{[z_k]_l} = \lambda_1$$

Per evitare di dover scegliere ogni volta una componente opportuna, si puo' usare σ_k ,

$$\tau_{z_k} z_k \underbrace{=}_{(1.1)} \sum_{i=1}^n \overline{\alpha_i} \alpha_i \overline{\lambda_i}^k \lambda_i^k = \sum_{i=1}^n |\alpha_i|^2 |\lambda_i|^{2k} = |\lambda_1|^{2k} \left(|\alpha_1|^2 + \sum_{i=2}^n |\alpha_i|^2 \left(\frac{|\lambda_i|}{|\lambda_1|}\right)^{2k} \right) \quad (3)$$

$$Az_k = z_{k+1} \Rightarrow \tau_{z_k} Az_k = \tau_{z_k} z_{k+1} \underbrace{=}_{(1.1)} \sum_{i=1}^n |\alpha_i|^2 \overline{\lambda_i}^k \lambda_i^{k+1} = \sum_{i=1}^n |\alpha_i|^2 |\lambda_i|^{2k} \lambda_i =$$

$$= |\lambda_1|^{2k} \lambda_1 \left(|\alpha_1|^2 + \sum_{i=2}^n |\alpha_i|^2 \left(\frac{|\lambda_i|}{|\lambda_1|}\right)^{2k} \frac{\lambda_i}{\lambda_1} \right) \quad (3.1)$$

$$\sigma'_k = \frac{\tau_{z_k} Az_k}{\tau_{z_k} z_k} \underbrace{=}_{(3), (3.1)} \lambda_1 \frac{|\alpha_1|^2 + \sum_{i=2}^n |\alpha_i|^2 \left(\frac{|\lambda_i|}{|\lambda_1|}\right)^{2k} \frac{\lambda_i}{\lambda_1}}{|\alpha_1|^2 + \sum_{i=2}^n |\alpha_i|^2 \left(\frac{|\lambda_i|}{|\lambda_1|}\right)^{2k}} \Rightarrow \lim_{k \rightarrow \infty} \sigma'_k = \lambda_1$$

dove con $\bar{\alpha}_i$ abbiamo indicato il complesso coniugato di α_i .
Poiche'

$$[z_k]_l = \lambda_1^k \left(\alpha_1 u_{1l} + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k u_{il} \right) \Rightarrow$$

$$\Rightarrow \lim_{k \rightarrow \infty} |[z_k]_l| = \begin{cases} +\infty & |\lambda_1| > 1 \\ |\alpha_i u_{il}| & \lambda_1 = 1 \\ 0 & 0 \leq |\lambda_1| < 1 \end{cases}$$

nell'implementazione macchina rischiamo di incorrere in errori di overflow o di underflow. Lo stesso problema si riscontra in $\tau_{z_k} z_k$ e in $\tau_{z_k} A z_k$, perche' figura $|\lambda_1|^{2k}$. Per questo motivo si usano gli x_k, y_k della tesi, in modo tale da avere gli y_k della successione con $\|y_k\| = 1$. I calcoli svolti fin adesso non variano, infatti, si vede facilmente che

$$y_k = \frac{z_k}{\|z_k\|}, \text{ infatti, per induzione:}$$

$$y_0 = \frac{x_0}{\|x_0\|} = \frac{z_0}{\|z_0\|}$$

$$x_{k+1} = A y_k \underbrace{=}_{\text{Hp induttiva}} A \frac{z_k}{\|z_k\|} = \frac{z_{k+1}}{\|z_k\|} \quad (*)$$

$$y_{k+1} = \frac{x_{k+1}}{\|x_{k+1}\|} \underbrace{=}_{(*)} \frac{z_{k+1}}{\|z_k\| \frac{\|z_{k+1}\|}{\|z_k\|}} = \frac{z_{k+1}}{\|z_{k+1}\|}$$

$$x_{k+1} = A y_k = \frac{z_{k+1}}{\|z_k\|}$$

$$\sigma_{k+1} = \frac{{}^t x_{k+1} A x_{k+1}}{{}^t x_{k+1} x_{k+1}} = \frac{{}^t z_{k+1} A \frac{z_{k+1}}{\|z_k\|}}{{}^t z_{k+1} \frac{z_{k+1}}{\|z_k\|}} = \frac{{}^t z_{k+1} A z_{k+1}}{{}^t z_{k+1} z_{k+1}} = \sigma'_{k+1}$$

Stavola i calcoli sono stabili, infatti,

$$\tau_{y_k} y_k = \frac{\tau_{z_k} z_k}{\|z_k\| \|z_k\|} = \frac{z_k \cdot z_k}{\|z_k\|^2} = 1$$

$$\tau_{y_k} A y_k = \lambda_1 \frac{(\dots)}{(\dots)} \quad [\text{non compare piu' } |\lambda_1|^{2k}]$$

(1) Dim. $\frac{\tau_{y_k} x_{k+1}}{\tau_{y_k} y_k} = \sigma_k$
Infine abbiamo

$$\frac{\tau_{y_k} x_{k+1}}{\tau_{y_k} y_k} = \frac{\frac{\tau_{x_k}}{\|x_k\|} x_{k+1}}{\frac{\tau_{x_k}}{\|x_k\|} \frac{x_k}{\|x_k\|}} = \frac{\tau_{x_k} x_{k+1}}{\tau_{x_k} \frac{x_k}{\|x_k\|}} =$$

$$= \frac{\tau_{x_k} A y_k}{\tau_{x_k} \frac{x_k}{\|x_k\|}} = \frac{\tau_{x_k} A \frac{\tau_{x_k}}{\|x_k\|}}{\tau_{x_k} \frac{x_k}{\|x_k\|}} = \frac{\tau_{x_k} A x_k}{\tau_{x_k} x_k} = \sigma_k$$

(2) Calcolo di λ_h
Abbiamo

$$x'_0 = x_0 - \sum_{i=1}^{h-1} \alpha_i u_i = \sum_{i=h}^n \alpha_i u_i$$

In pratica, da x_0 abbiamo eliminato tutti i termini precedenti a $\alpha_h u_h$. Adesso, il λ_i dominante e' λ_h . Ripetendo in modo analogo i ragionamenti di prima, si ha la tesi.

(3) Criterio di arresto

Il criterio di arresto, si basa sulle seguenti osservazioni:

$$\begin{aligned}
Au_1 &= \lambda_1 u_1 \Leftrightarrow (A - \lambda_1 I)u_1 = 0 \\
\lim_{k \rightarrow \infty} y_k &= u_1 \\
\lim_{k \rightarrow \infty} \sigma_k &= \lambda_1 \\
\lim_{k \rightarrow \infty} (A - \sigma_k I)y_k &= \underbrace{0}_{\text{vettore nullo}} \\
\lim_{k \rightarrow \infty} \|(A - \sigma_k I)y_k\| &= 0
\end{aligned}$$

□

Proposition 5.16. *Metodo delle potenze inverse per il calcolo dell'autovalore minimo. Siano $\lambda_1, \dots, \lambda_n$ tutti gli autovalori di $A \in \mathbb{C}^{n \times n}$, matrice diagonalizzabile. Sia*

$$\begin{aligned}
x_0 &\in \mathbb{C}^{n,1}, \quad x_0 \neq 0 \\
\begin{cases} y_k = \frac{x_k}{\|x_k\|} \\ x_{k+1} = A^{-1}y_k \end{cases} &\quad \forall k = 0, 1, 2, \dots \\
\sigma_k &= \frac{{}^t y_k A^{-1} y_k}{{}^t y_k y_k}
\end{aligned}$$

Supponiamo che

$$\begin{aligned}
&\exists A^{-1} \\
&|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{n-1}| > |\lambda_n|
\end{aligned}$$

Allora si ha:

$$\lim_{k \rightarrow \infty} \sigma_k = \frac{1}{\lambda_n}$$

Al posto di calcolare A^{-1} , per poi calcolare i termini della successione:

$$x_{k+1} = A^{-1}y_k$$

conviene calcolare i termini risolvendo di volta in volta il sistema

$$Ax_{k+1} = y_k$$

nell'incognita x_{k+1} .

Poiche' i singoli sistemi hanno la stessa matrice dei coefficienti, si puo' usare la prop. [5.9,pg.26].

Proof: Basta osservare che

$$\lambda_i \text{ autovalore di } A \Leftrightarrow \frac{1}{\lambda_i} \text{ autovalore di } A^{-1}$$

$$\frac{1}{|\lambda_1|} \leq \frac{1}{|\lambda_2|} \leq \dots \leq \frac{1}{|\lambda_{n-1}|} < \frac{1}{|\lambda_n|}$$

$$\lambda_{\max}(A^{-1}) = \frac{1}{\lambda_{\min}(A)} \Leftrightarrow \lambda_{\max}(A^{-1}) = \frac{1}{\lambda_n}$$

(vedi la dimostrazione di [4.10,pg.12]) e applicare la prop. [5.15,pg.29].

□

Proposition 5.17. Nelle ipotesi della prop [5.16,pg.33], preso un $\mu_0 \in \mathbb{C}$ e ponendo

$$A_{\mu_k} = A - \mu_k I$$

$\sigma^{(k)}$ approssimazione di σ_i di A_{μ_k} calcolato con la prop. [5.16,pg.33], con errore ε

$$\mu_{k+1} = \mu_k + \frac{1}{\sigma^{(k)}}$$

si ha

$$\lim_{k \rightarrow \infty} \mu_k = \lambda_\mu$$

dove λ_μ e' l'autovalore di A piu' vicino a μ , ovvero

$$|\lambda_\mu - \mu| \leq |\lambda_i - \mu| \quad \forall i = 1, \dots, n$$

Proof: Si ha

λ autovalore di $A \Leftrightarrow Au = \lambda u \Leftrightarrow Au - \mu u = \lambda u - \mu u \Leftrightarrow (A - \mu I)u = (\lambda - \mu)u \Leftrightarrow A_\mu u = (\lambda - \mu)u \Leftrightarrow \lambda - \mu$ autovalore

$\Rightarrow$ Gli autovalori di A_μ sono $\{\lambda_i - \mu \mid i = 1, \dots, n\}$ (1)

Quindi applicando la prop [5.16,pg.33] a A_μ , abbiamo che

$$\sigma^{(0)} \approx \frac{1}{\lambda_{\min}(A_\mu)} \underbrace{=}_{(1)} \frac{1}{\lambda_t - \mu}$$

dove $|\lambda_t - \mu| \leq |\lambda_i - \mu| \quad \forall i = 1, \dots, n$

$$\sigma^{(0)} \approx \frac{1}{\lambda_t - \mu} \Leftrightarrow \frac{1}{\sigma^{(0)}} \approx \lambda_t - \mu \Leftrightarrow \frac{1}{\sigma^{(0)}} + \mu \approx \lambda_t$$

Ponendo $\mu_1 = \frac{1}{\sigma^{(0)}} + \mu$, e reiterando, otteniamo

$$\frac{1}{\sigma^{(1)}} + \mu_1 \approx \lambda_t$$

che e' una approssimazione migliore di λ_t . □

6 Interpolazione e approssimazione

Definition 6.1. Supponiamo di avere gli $n+1$ punti $(x_i, y_i) \in \mathbb{R}^2 \quad \forall i = 0, 1, \dots, n$. L'interpolazione consiste nel trovare una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ tale che

$$f(x_i) = y_i \quad \forall i = 0, 1, \dots, n$$

Interpolazione polinomiale Si richiede che $f(x)$ sia un polinomio in $\mathbb{R}[x]$ di grado minimo.

Interpolazione razionale Si richiede che $f(x)$ sia una funzione razionale, ovvero che si del tipo

$$\frac{g(x)}{h(x)}, \quad g(x), h(x) \in \mathbb{R}[x]$$

Interpolazione trigonometrica Si richiede che $f(x)$ sia una funzione di periodo L . Si ricerca una soluzione tra le funzioni del tipo

$$\frac{a_0}{2} + \sum_{k=1}^n \left(b_k \sin \left(k \frac{2\pi}{L} x \right) + a_k \cos \left(k \frac{2\pi}{L} x \right) \right)$$

6.1 Interpolazione polinomiale

Proposition 6.2. Sia $f(x) = a_0 + a_1x + \dots + a_nx^n$. Un algoritmo per computare $f(x)$ in maniera efficace e' il seguente:

$$\begin{aligned} t &:= a_n \\ \forall i &= 1, \dots, n \\ t &:= a_{n-i} + tx \end{aligned}$$

Il costo dell'algoritmo e' quindi

n moltiplicazioni, n addizioni

Proposition 6.3.

$$\exists! p(x) \in \mathbb{R}[x]_n : p(x_i) = y_i \quad \forall i = 0, 1, \dots, n \Leftrightarrow x_i \neq x_j \quad \forall i > j$$

dove $\mathbb{R}[x]_n$ e' lo spazio vettoriale dei polinomi di grado $\leq n$ a coefficienti reali.

Proof: Sia $p(x)$ il generico polinomio di grado $\leq n$ incognita:

$$\begin{aligned} p(x) &= \sum_{j=0}^n a_j x^j \\ p(x_i) &= y_i \quad \forall i = 0, \dots, n \Leftrightarrow \sum_{j=0}^n a_j x_i^j = y_i \quad \forall i = 0, \dots, n \end{aligned}$$

Quest'ultimo e' un sistema lineare nelle incognite $a_0, \dots, a_n$, in $n+1$ equazioni. La matrice dei coefficienti e'

$$V = \begin{bmatrix} x_0^0 & x_0^1 & \dots & x_0^n \\ x_1^0 & x_1^1 & \dots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ x_n^0 & x_n^1 & \dots & x_n^n \end{bmatrix}$$

Il suo determinante e' il noto determinante di Vandermonde:

$$|V| = \prod_{i>j} (x_i - x_j)$$

Il sistema ammette un'unica soluzione $\Leftrightarrow |V| \neq 0$, ovvero

$$|V| \neq 0 \Leftrightarrow x_i - x_j \neq 0 \quad \forall i > j$$

□

Theorem 6.4. Forma di Lagrange.

Sia

$$\lambda_j(x) = \prod_{0 \leq i \leq n, i \neq j} \frac{x - x_i}{x_j - x_i}$$

allora la soluzione dell'interpolazione e'

$$p(x) = \sum_{j=0}^n y_j \lambda_j(x)$$

Proof: Nella dimostrazione [6.3,pg.35], abbiamo visto che trovare $p(x)$ equivale a risolvere il sistema

$$Va = y$$

$$\text{con } a = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}, \quad y = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}$$

PROOF SKETCH: Poiche' $p(x) \in \mathbb{R}[x]$ e $\mathbb{R}[x]_n$ e' uno spazio vettoriale, possiamo considerare una base $\{\lambda_0(x), \lambda_1(x), \dots, \lambda_n(x)\}$ in modo tale che la nuova matrice dei coefficienti V' sia I e che $\deg \lambda_j(x) \leq n$ Ovvero:

$$p(x) = \sum_{j=0}^n a'_j \lambda_j(x) \quad (1)$$

$$p(x_i) = \sum_{j=0}^n a'_j \lambda_j(x_i) \quad \forall i = 0, 1, \dots, n \quad (1.1)$$

$$(V')_{ij} = \lambda_j(x_i)$$

$$V = I \Leftrightarrow (V)_{ij} = \delta_{ij} \Leftrightarrow \lambda_j(x_i) = \delta_{ij}$$

Ricaviamoci i $\lambda_j(x)$:

$$\lambda_j(x_i) = \delta_{ij} \quad \forall i \Rightarrow \lambda_j(x_i) = 0 \quad \forall i \neq j$$

$$\begin{cases} \lambda_j(x_i) = 0 \quad \forall i = 0, 1, \dots, j-1, j+1, \dots, n \\ \deg \lambda_j(x) \leq n \end{cases} \quad \underbrace{\Rightarrow}_{\text{Ruffini}} \lambda_j(x) = c \prod_{i \neq j} x - x_i$$

$$\lambda_j(x_j) = 1 \Leftrightarrow c \prod_{i \neq j} x_j - x_i = 1 \Leftrightarrow c = \frac{1}{\prod_{i \neq j} x_j - x_i}$$

$$\Rightarrow \lambda_j(x) = \frac{\prod_{i \neq j} x - x_i}{\prod_{i \neq j} x_j - x_i} = \prod_{0 \leq i \leq n, i \neq j} \frac{x - x_i}{x_j - x_i}$$

Quindi

$$p(x_i) = y_i \quad \underbrace{\Leftrightarrow}_{(1.1)} p(x_i) = \sum_{j=0}^n a'_j \delta_{ij} = y_i \quad \forall i = 0, 1, \dots, n \Leftrightarrow \begin{cases} a'_0 = y_0 \\ a'_1 = y_1 \\ \vdots \\ a'_n = y_n \end{cases}$$

$$p(x) \quad \underbrace{=}_{(1)} \sum_{j=0}^n a'_j \lambda_j(x) = \sum_{j=0}^n y_j \lambda_j(x)$$

□

Proposition 6.5. Nel contesto del thm [6.4,pg.35] si ha

$$\sum_{j=0}^n \lambda_j(x) (x_j - x)^k = \begin{cases} 1 & k = 0 \\ 0 & k > 0 \end{cases} \quad \forall x \in \mathbb{R}$$

Proof: Sia $f_k(x)$ il polinomio della tesi.

$$f_0(x) = \sum_{j=0}^n \lambda_j(x)$$

quindi $f_0(x)$ interpola i punti $(x_j, 1)$ $j = 0, \dots, n$, quindi per l'unicita' dell'interpolazione (vedi [6.3,pg.35]), si ha $f_0 \equiv 1$.

Se $k > 0$, si ha

$$\lambda_j(x_i)(x_j - x_i)^k = \delta_{ij}(x_j - x_i)^k = \begin{cases} 0(x_j - x_i)^k & i \neq j \\ 1 \cdot 0 & i = j \end{cases} = 0 \Rightarrow f_k(x_i) = 0 \quad \forall i = 0, \dots, n$$

quindi $f_k(x)$ interpola i punti $(x_i, 0)$. Allora, per l'unicita' dell'interpolazione $f_k \equiv 0$. $\square$

Definition 6.6. Sia $f \in C^0([a, b])$, $x_i \in [a, b] \quad \forall i = 0, 1, \dots, n$ definiamo la *differenza divisa*:

$$\begin{aligned} f[x_0] &= f(x_0) \quad \text{ovvero la funzione } f \text{ calcolata in } x_0 \\ f[x_0, \dots, x_j] &= \frac{f[x_1, \dots, x_j] - f[x_0, \dots, x_{j-1}]}{x_j - x_0}, \quad \text{per } j > 0 \end{aligned}$$

Proposition 6.7. *Il seguente e' un algoritmo che calcola $f[x_0, \dots, x_n]$ in $\frac{n(n+1)}{2}$ flops:*

$$\begin{aligned} a_i &:= f(x_i) \quad \forall i = 0, \dots, n \\ \forall k &= n-1, \dots, 1, 0 \\ \forall i &= 0, \dots, k \\ a_i &:= \frac{a_{i+1} - a_i}{x_{i+n-k} - x_i} \end{aligned}$$

Il risultato e' in a_0

si puo' visualizzare l'algoritmo cosi':

$$\begin{array}{llll} f[x_0] &= f(x_0) & & \\ & & f[x_0, x_1] & \\ f[x_1] &= f(x_1) & & f[x_0, x_1, x_2] \\ & & f[x_1, x_2] & f[x_0, x_1, x_2, x_3] \\ f[x_2] &= f(x_2) & & f[x_1, x_2, x_3] \\ & & f[x_2, x_3] & \\ f[x_3] &= f(x_3) & & \end{array}$$

Theorem 6.8. *Sia $f \in C^0([a, b])$, $x_i \in [a, b] \quad \forall i = 0, 1, \dots, n$. Il polinomio di Newton che interpola i punti $(x_i, f(x_i))$ e'*

$$\begin{aligned} p_n(x) &= \sum_{i=0}^n f[x_0, \dots, x_i] w_i(x) \\ p_n(x_i) &= f(x_i) \quad \forall i = 0, 1, \dots, n & \text{dove} \\ w_0(x) &= 1 \\ w_j(x) &= \prod_{i=0}^{j-1} (x - x_i) \quad \forall j > 0 \end{aligned}$$

$p_n(x)$ ha il vantaggio di poter essere facilmente aggiornato per l'aggiunta di nuovi punti. Infatti, dato x_{n+1} , basta calcolare

$$p_{n+1}(x) = p_n(x) + f[x_0, \dots, x_{n+1}] w_{n+1}(x)$$

Inoltre, vale la seguente proposizione:

$$f(x) = p_n(x) + R_n(x) \quad \forall x \in [a, b]$$

dove

$$R_n(x) = f[x_0, \dots, x_n, x] w_{n+1}(x)$$

$$w_{n+1}(x) = \prod_{i=0}^n (x - x_i)$$

Infine, per calcolare $p_n(x)$ in x , si puo' usare il seguente algoritmo che si basa sulla prop [6.2,pg.35]:

$$t := f[x_0, \dots, x_n]$$

$$\forall i = 1, \dots, n$$

$$t := f[x_0, \dots, x_{n-i}] + t \cdot (x - x_{n-i})$$

Dove $f[x_0, \dots, x_{n-i}]$ si puo' calcolare con la prop [6.7,pg.37] (nota⁶)

Proof: Poniamo $p_0(x) = f(x_0)$, $p_{-1}(x) = 0$

(1) Dim. che

$$\exists A_1, \dots, A_n \in \mathbb{R} : p_k(x) = p_{k-1}(x) + A_k w_k(x) = \sum_{i=0}^k A_i w_i(x) \quad \forall k = 0, \dots, n$$

dove $p_k(x)$ e' il polinomio interpolatore di $(x_i, f(x_i)) \quad \forall i = 0, \dots, k$.

Per Hp abbiamo:

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

Usando un polinomio incognita $g(x)$ di grado $\leq n+1$:

$$p_{n+1}(x) = p_n(x) + g(x)$$

$$p_{n+1}(x_i) = f(x_i) = p_n(x_i) \quad \forall i = 0, \dots, n \Rightarrow g(x_i) = 0 \quad \forall i = 0, \dots, n \quad (1)$$

$$\begin{cases} \deg p_{n+1}(x) \leq n+1 \\ \deg p_n(x) \leq n \end{cases} \Rightarrow \deg g(x) \leq n+1 \quad (2)$$

$$(1), (2) \xRightarrow{\text{Ruffini}} g(x) = A_{n+1} \prod_{i=0}^n (x - x_i) = A_{n+1} w_{n+1}(x)$$

$$p_{n+1}(x) = p_n(x) + A_{n+1} w_{n+1}(x) \quad (3)$$

⁶in realta', non serve calcolare a ogni passo $f[x_0, \dots, x_{n-i}]$ usando [6.7,pg.37], perche' alcuni calcoli vengono ripetuti inutilmente. Si puo' quindi pensare a un algoritmo piu' ottimizzato.

In particolare abbiamo:

$$p_0(x) = f(x_0) = \sum_{i=0}^0 A_i w_i(x)$$

$$p_1(x) = f(x_0) + A_1(x - x_0) = \sum_{i=0}^1 A_i w_i(x)$$

$$p_2(x) = f(x_0) + A_1(x - x_0) + A_2(x - x_0)(x - x_1) = \sum_{i=0}^2 A_i w_i(x)$$

...

$$p_n(x) = \sum_{i=0}^n A_i w_i(x)$$

Adesso resta da determinare gli A_i .

(2) Dim. che vale la seguente implicazione:

$$p_n(x) \text{ e' il polinomio interpolatore di } (x_i, f(x_i)) \quad \forall i = 0, \dots, n \Rightarrow p_n(x) = \sum_{i=0}^n f[x_0, \dots, x_i] w_i(x)$$

Procediamo per induzione su n .

(2.1) $n = 0$

$$p_0(x) = f(x_0) = f[x_0]$$

(2.2) $n = 1$

$$p_1(x) \underbrace{=}_{\text{passo 1}} p_0(x) + A_1(x - x_0) \underbrace{=}_{\text{passo 2.1}} f[x_0] + A_1(x - x_0)$$

Imponendo che $p_1(x_1) = f(x_1)$, abbiamo

$$A_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = F[x_0, x_1]$$

(2.3) $n - 1 \Rightarrow n$

Sia $h_{n-1}(x)$ il polinomio d'interpolazione relativo ai punti $x_1, \dots, x_n$, e $g_{n-1}(x)$ quello relativo ai punti $x_0, \dots, x_{n-1}$, ovvero

$$h_{n-1}(x_i) = f(x_i) \quad \forall i = 1, \dots, n \quad (4)$$

$$g_{n-1}(x_i) = f(x_i) \quad \forall i = 0, \dots, n-1 \quad (4.1)$$

Si ha

$$p_n(x) = \frac{x - x_0}{x_n - x_0} h_{n-1}(x) + \frac{x_n - x}{x_n - x_0} g_{n-1}(x) \quad (5)$$

infatti,

$$x = x_0 \Rightarrow \frac{x_0 - x_0}{x_n - x_0} h_{n-1}(x_0) + \frac{x_n - x_0}{x_n - x_0} g_{n-1}(x_0) = g_{n-1}(x_0) \underbrace{=}_{(4.1)} f(x_0) = p_n(x_0)$$

$$x = x_n \Rightarrow \frac{x_n - x_0}{x_n - x_0} h_{n-1}(x_n) + \frac{x_n - x_n}{x_n - x_0} g_{n-1}(x_n) = h_{n-1}(x_n) \underbrace{=}_{(4)} f(x_n) = p_n(x_n)$$

$$x = x_i, \quad 0 < i < n \Rightarrow \frac{x_i - x_0}{x_n - x_0} h_{n-1}(x_i) + \frac{x_n - x_i}{x_n - x_0} g_{n-1}(x_i) =$$

$$\underbrace{=}_{h_{n-1}(x_i)=f(x_i)=g_{n-1}(x_i)} \left(\frac{x_i - x_0}{x_n - x_0} + \frac{x_n - x_i}{x_n - x_0} \right) f(x_i) = f(x_i) = p_n(x_i)$$

e quindi per l'unicita' di $p_n(x)$ (vedi [6.3,pg.35]) si ha l'uguaglianza polinomiale nella (5).

Applicando l'ipotesi induttiva su h_{n-1} e g_{n-1} abbiamo⁷

$$\begin{aligned} h_{n-1}(x) &= f[x_1] + f[x_1, x_2](x - x_1) + f[x_1, x_2, x_3](x - x_1)(x - x_2) + \dots + \\ &\quad + f[x_1, x_2, \dots, x_n](x - x_1)(x - x_2) \dots (x - x_{n-1}) \\ g_{n-1}(x) &= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \dots + \\ &\quad + f[x_0, x_1, \dots, x_{n-1}](x - x_0)(x - x_1) \dots (x - x_{n-2}) \end{aligned}$$

Indicando con $M(h_{n-1})$ il coefficiente dell'indeterminata grado massimo di h_{n-1} , abbiamo
 $M(h_{n-1}) = f[x_1, x_2, \dots, x_n]$

$$M(g_{n-1}) = f[x_0, x_1, \dots, x_{n-1}]$$

$$M(p_n) \underset{(5)}{=} \frac{M(h_{n-1}) - M(g_{n-1})}{x_n - x_0} = \frac{f[x_1, x_2, \dots, x_n] - f[x_0, x_1, \dots, x_{n-1}]}{x_n - x_0} \underset{\text{per def.}}{=} f[x_0, \dots, x_n] \quad (5.1)$$

D'altra parte,

$$\begin{cases} p_n(x) \underset{(3)}{=} p_{n-1}(x) + A_n w_n(x) \\ \deg p_{n-1}(x) \leq n-1 \\ \deg A_n w_n(x) \leq n \\ w_n(x) \text{ e' monico} \end{cases} \Rightarrow M(p_n) = M(A_n w_n(x)) = A_n \underset{(5.1)}{=} f[x_0, \dots, x_n]$$

$$p_n(x) = p_{n-1}(x) + f[x_0, \dots, x_n] w_n(x) =$$

$$\underset{\text{Hp induttiva su } p_{n-1}}{=} \sum_{i=0}^{n-1} f[x_0, \dots, x_i] w_i(x) + f[x_0, \dots, x_n] w_n(x) = \sum_{i=0}^n f[x_0, \dots, x_i] w_i(x)$$

(3) Dim. che $f(x) = p_n(x) + R_n(x) \quad \forall x \in [a, b]$

Fissiamo $x \in [a, b]$. Se consideriamo $(x_{n+1}, f(x_{n+1})) \equiv (x, f(x))$ come un nuovo punto di interpolazione, abbiamo

$$f(x) = p_{n+1}(x) \underset{\text{passo 2}}{=} p_n(x) + f[x_1, \dots, x] w_{n+1}(x) = p_n(x) + R_n(x)$$

□

Proposition 6.9. *Nel contesto di [6.8, pg. 37], si ha*

$$\exists f^{(n+1)}(x) \text{ in } [a, b] \Rightarrow \exists \xi \in [a, b] : R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x)$$

dove $f^{(n+1)}(x)$ e' la derivata $n+1$ -esima di f .

Proposition 6.10. *Sia*

$$\begin{aligned} t_n(x) &= \cos n\varphi(x) : [-1, 1] \longrightarrow \mathbb{R} \\ \varphi(x) &= \arccos x : [-1, 1] \longrightarrow [0, \pi] \end{aligned}$$

$t_n(x)$ si puo' esprimere come un polinomio, chiamato polinomio di Chebyshev:

$$\begin{aligned} t_0(x) &= 1 \\ t_1(x) &= x \\ t_n(x) &= 2t_{n-1}(x)x - t_{n-2}(x) \quad \forall n > 1 \end{aligned}$$

⁷se volessimo applicare l'Hp induttiva formalmente, ad esempio su h_{n-1} , basterebbe porre $y_0 = x_1, y_1 = x_2, \dots, y_{n-1} = x_n$

Valgono le seguenti proprietà:

$$\begin{aligned}
\deg t_n(x) &= n \quad \forall n \geq 0 \\
c_n &= 2^{n-1} \quad \text{coeff. di grado massimo di } t_n(x) \quad \forall n \geq 1 \\
\alpha_k &= \cos\left(\frac{\pi}{2n} + k\frac{\pi}{n}\right) \quad k = 0, \dots, n-1 \quad \text{sono tutte le radici di } t_n(x) \\
M_k &= \cos\left(k\frac{2\pi}{n}\right) \quad k = 0, \dots, n \quad \text{sono tutti i punti di massimo} \\
m_k &= \cos\left(k\frac{2\pi}{n} + \frac{\pi}{n}\right) \quad k = 0, \dots, n \quad \text{sono tutti i punti di minimo}
\end{aligned}$$

Proof:

$$t_{n+1}(x) = \cos((n+1)\varphi(x)) = \cos(n\varphi(x) + \varphi(x)) = \cos(n\varphi(x))\cos(\varphi(x)) - \sin(n\varphi(x))\sin(\varphi(x))$$

$$t_{n-1}(x) = \cos((n-1)\varphi(x)) = \cos(n\varphi(x) - \varphi(x)) = \cos(n\varphi(x))\cos(\varphi(x)) + \sin(n\varphi(x))\sin(\varphi(x))$$

Sommando membro a membro:

$$t_{n+1}(x) + t_{n-1}(x) = 2\cos(n\varphi(x))\cos(\varphi(x)) = 2t_n(x)x \Leftrightarrow t_{n+1}(x) = 2t_n(x)x - t_{n-1}(x)$$

$$t_n(x) = 2t_{n-1}(x)x - t_{n-2}(x)$$

(1) $\deg t_n(x) = n$

Per induzione forte su n : I casi $n = 0, 1$ sono veri.

$$\begin{aligned}
\deg t_n(x) &= \deg [2t_{n-1}(x)x - t_{n-2}(x)] \\
\deg t_{n-1}(x)x &= \deg t_{n-1}(x) + 1 \quad \underbrace{\hspace{1cm}}_{\text{Hp induttiva}} \quad (n-1) + 1 = n
\end{aligned}$$

$$\deg t_{n-2}(x) \quad \underbrace{\hspace{1cm}}_{\text{Hp induttiva}} \quad n-2$$

$$\Rightarrow \deg t_n(x) = n$$

(2) $c_n = 2^{n-1}, n \geq 1$

$$t_1(x) = x \Rightarrow c_1 = 1 = 2^{1-1}$$

$$\begin{cases} t_n(x) = 2t_{n-1}(x)x - t_{n-2}(x) \\ \deg 2t_{n-1}(x)x = n \\ \deg t_{n-2} = n-2 \end{cases} \Rightarrow c(t_n) = c(2t_{n-1}(x)x) = 2c(t_{n-1}(x)x) = 2c(t_{n-1}) \quad \underbrace{\hspace{1cm}}_{\text{Hp induttiva}} \quad 22^{n-2} = 2^{n-1}$$

dove con $c(t_n)$ abbiamo indicato il termine di grado massimo di $t_n(x)$.

Infine, per ricercare le radici, i massimi e i minimi, basta cercare i punti in cui la funzione $\cos n\varphi$ assume rispettivamente i valori 0, 1, -1, e ricavarsi la x sapendo che $x = \cos(\varphi)$. $\square$

Proposition 6.11. *Nel contesto della prop [6.10, pg.40] si ha*

$$\frac{1}{2^{n-1}} = \left\| \frac{1}{2^{n-1}} t_n(x) \right\|_{[-1,1]} = \min \left\{ \|p\|_{[-1,1]} \mid p \in \mathbb{R}[x]_n, \text{ } p \text{ monico} \right\}, \quad \text{con } n \geq 1$$

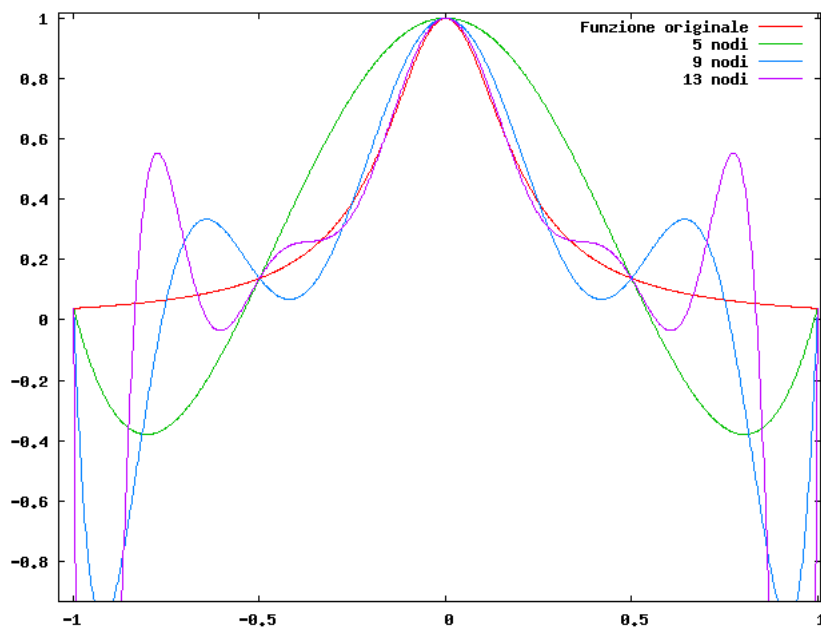
Nota: $\frac{1}{2^{n-1}} t_n(x)$ e' un polinomio monico.

Example 6.12. Fenomeno di Runge.

Consideriamo la funzione

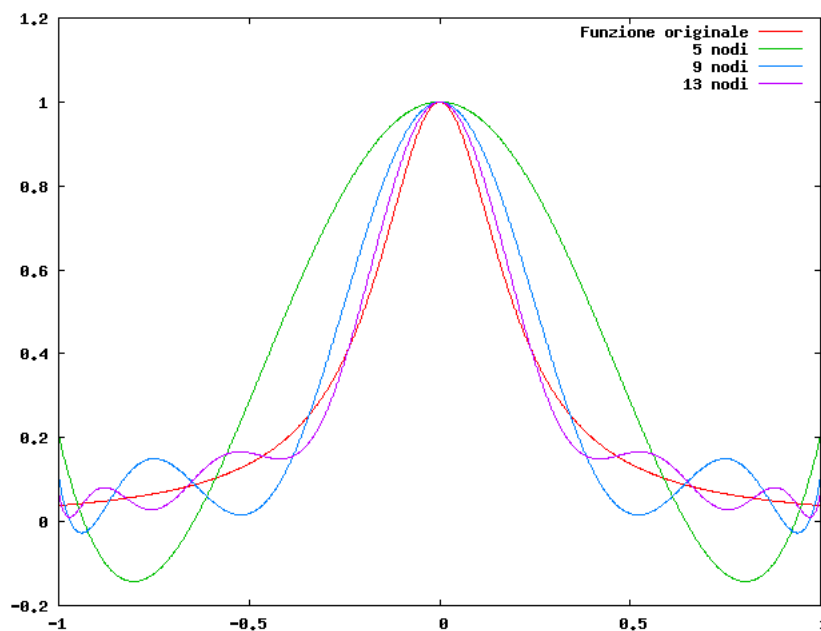
$$f(x) = \frac{1}{1 + 25x^2} : [-1, 1] \longrightarrow \mathbb{R}$$

Se la interpoliamo in $n + 1$ punti distribuiti uniformemente in $[-1, 1]$, otteniamo una pessima approssimazione di $f(x)$, ovvero esistono alcuni punti $x \in [-1, 1]$ dove $|f(x) - p_n(x)|$ e' molto alto. Al crescere di n la situazione peggiora: si dimostra che per $n \rightarrow \infty$ esistono degli $x \in [-1, 1]$ t.c. $|f(x) - p_n(x)| \rightarrow \infty$. Vedi figura:



$f(x)$ e' la linea in rosso.

Se invece scegliamo le radici di Chebyshev come punti per l'interpolazione, abbiamo:



La miglioria e' dovuta al fatto che le radici di Chebyshev sono piu' distribuite verso gli estremi dell'intervallo $[-1, 1]$.

I grafici sono stati prodotti con il seguente script octave:

```
clf
hold on
xf=[-1:0.002:1]; yf=1./(1+25*xf.^2);
plot(xf,yf,';Funzione originale;')

nodi=[5:4:16];
steps=2./(nodi-1);

function yi=polyinterp(x,y,xi)
    n = length(x);
    M = zeros(n);
    for i=1:n
        M(:,i) = (x.^(i-1))';
    endfor
    a = M \ y';
    g = ones(1,length(xi)) * a(1);
    for i=2:n
        g = g + xi.^(i-1)*a(i);
    endfor
    yi=g;
endfunction

for i=1:length(nodi)
    xp=[-1:steps(i):1];
    yp=1./(1+25*xp.^2);
    Y=polyinterp(xp,yp,xf);
    plot(xf,Y, sprintf(";%d nodi;",nodi(i)))
end;

pause

%Chebyshev
clf
plot(xf,yf,';Funzione originale;')
for i=1:(length(nodi))
    n=nodi(i)
    k=[0:n]
    xp=cos(pi.*(2.*k+1)./(2*n));
    yp=1./(1+25*xp.^2);
    Y=polyinterp(xp,yp,xf);
    plot(xf,Y, sprintf(";%d nodi;",n))
end;
```

Proposition 6.13. *Siano fissati $n + 1$ punti $x_0, x_1, \dots, x_n \in [a, b]$. Sia definita la seguente*

funzione:

$$L_n : C^0([a, b]) \longrightarrow \mathbb{R}[x]_n$$

$L_n(f)$ e' il polinomio che interpola f nei punti $(x_i, f(x_i)) \quad \forall i = 0, \dots, n$

che e' ben definita per la prop [6.3,pg.35].

Si ha:

L_n e' una applicazione lineare dello spazio vettoriale $C^0([a, b])$

Proof:

$$L_n(\alpha f + \beta g)(x_i) = (\alpha f + \beta g)(x_i) = \alpha f(x_i) + \beta g(x_i) = \alpha L_n(f)(x_i) + \beta L_n(g)(x_i) \quad \forall i = 0, 1, \dots, n$$

$\Rightarrow L_n(f)(x) + \beta L_n(g)(x)$ e' un polinomio interpolatore di $\alpha f + \beta g$ in $x_i \quad \forall i = 0, 1, \dots, n$

Quindi, per l'unicita' del polinomio interpolatore (vedi [6.3,pg.35]) si ha

$$L_n(\alpha f + \beta g) = \alpha L_n(f) + \beta L_n(g)$$

□

7 Approssimazione polinomiale

Definition 7.1. L'approssimazione polinomiale consiste nel trovare un polinomio $p(x)$ che approssimi una funzione $f(x) \in C^0([a, b])$.

Proposition 7.2. La funzione

$$\begin{aligned} \|\cdot\| : C^0([a, b]) &\longrightarrow \mathbb{R}_+ \\ \|f\| &= \max_{x \in [a, b]} |f(x)| \end{aligned}$$

e' una norma nello spazio vettoriale $C^0([a, b])$.

Theorem 7.3. di Weierstrass.

$$(\mathbb{R}[x], \|\cdot\|) \text{ e' denso in } (C^0([a, b]), \|\cdot\|)$$

dove $\|\cdot\|$ e' la norma definita in [7.2,pg.44], e $(\mathbb{R}[x]_n, \|\cdot\|)$, $(C^0([a, b]), \|\cdot\|)$ sono spazi metrici indotti da essa.

Theorem 7.4. di Weierstrass II.

Dato $n \geq 0$ e $f \in C^0([a, b])$, si ha

$$\exists \varphi_n \in \mathbb{R}[x]_n : \forall \varphi \in \mathbb{R}[x]_n \quad \|f - \varphi_n\| \leq \|f - \varphi\|$$

Inoltre,

$$\lim_{n \rightarrow +\infty} \varphi_n \stackrel{u.}{=} f \text{ in } [a, b]$$

dove si e' intesa una convergenza uniforme.

Definition 7.5. Definiamo il polinomio di Bernstein:

$$B_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}$$

Proposition 7.6. Sia $B_{n,k}(x)$ il polinomio di Bernstein, si ha

1. $\sum_{k=0}^n B_{n,k}(x) = 1$
2. $\{B_{n,0}, B_{n,1}, \dots, B_{n,n}\}$ e' una base di $K[x]_n$

Proof:

(1) Dim 1.

$$1 = x + 1 - x$$

$$1 = (x + (1 - x))^n = \sum_{k=0}^n \binom{n}{k} x^k (1 - x)^{n-k} = \sum_{k=0}^n B_{n,k}(x)$$

□

Theorem 7.7. Data una funzione $f(x) \in C^0([0, 1])$, e posto

$$B_n(f)(x) = \sum_{k=0}^n f\left(\frac{k}{n}\right) B_{n,k}(x)$$

si ha

$$\lim_{n \rightarrow +\infty} B_n(f) \stackrel{u}{=} f \text{ in } [0, 1]$$

Theorem 7.8. di Faber.

Sia $V_k = \{x_1, \dots, x_k\}$, con $x_i \in [a, b] \quad \forall i, \quad x_i \neq x_j \quad \forall i \neq j$. Sia $f(x) \in C^0([a, b])$, $f_k(x)$ il polinomio interpolatore dei punti $(x_i, f(x_i)) \quad \forall i = 1, \dots, k$. Il teorema di Faber afferma che

$$\forall \{V_k\}_{k \in \mathbb{N}} \quad \exists f(x) \in C^0([a, b]) \text{ tale che } \lim_{k \rightarrow \infty} f_k(x) \neq f(x)$$

In altre parole non e' sempre possibile usare con efficacia un unico schema d'interpolazione per approssimare una qualunque funzione.

Proposition 7.9. Data $f(x) \in C^0([a, b])$ sia $L_n(f) \in \mathbb{R}[x]_n$ il polinomio che interpola f in $n+1$ punti $(x_i, f(x_i))$. Si ha:

$$\|f - L_n(f)\| \leq E_n(f)(1 + \Lambda_n)$$

$$\lim_{n \rightarrow +\infty} E_n(f) \stackrel{u}{=} 0 \text{ in } [a, b]$$

$$\lim_{n \rightarrow \infty} \Lambda_n = \infty$$

dove

$\| - \|$ e' la norma definita in [7.2,pg.44]

$E_n(f) = \|f - \varphi_n\|$, con φ_n polinomio dato dalla prop. [7.4,pg.44]

$\Lambda_n = \max_{x \in [a, b]} \sum_{i=0}^n |\lambda_i(x)|$ e' il numero di Lebesgue

dove $\lambda_j(x)$ sono dati da [6.4,pg.35] rispetto ai punti $x_0, x_1, \dots, x_n$

Proof:

$$\|f - L_n(f)\| = \|f - \varphi_n + \varphi_n - L_n(f)\| \leq \|f - \varphi_n\| + \|\varphi_n - L_n(f)\| \stackrel{[6.13, \text{pg.43}]}{=} \|f - \varphi_n\| + \|L_n(\varphi_n - f)\| =$$

$$= E_n(f) + \|L_n(\varphi_n - f)\| \quad (1)$$

$$\|L_n(\varphi_n - f)\| = \max_{x \in [a, b]} |L_n(\varphi_n - f)(x)| \stackrel{[6.4, \text{pg.35}]}{=} \max_{x \in [a, b]} \left| \sum_{i=0}^n (\varphi_n(x_i) - f(x_i)) \lambda_i(x) \right| \leq$$

$$\leq \max_{x \in [a, b]} \sum_{i=0}^n \left| (\varphi_n(x_i) - f(x_i)) \lambda_i(x) \right| \leq \max_{x \in [a, b]} \sum_{i=0}^n \|\varphi_n(x) - f(x)\| |\lambda_i(x)| =$$

$$= \max_{x \in [a, b]} \|\varphi_n(x) - f(x)\| \sum_{i=0}^n |\lambda_i(x)| = \|\varphi_n(x) - f(x)\| \max_{x \in [a, b]} \sum_{i=0}^n |\lambda_i(x)| = E_n(f) \Lambda_n$$

$$(1) \Rightarrow \|f - L_n(f)\| \leq E_n(f) + E_n(f) \Lambda_n = E_n(f)(1 + \Lambda_n)$$

Infine, per il thm [7.4, pg.44]

$$\lim_{n \rightarrow +\infty} \varphi_n \stackrel{u}{=} f \text{ in } [a, b] \Rightarrow \lim_{n \rightarrow +\infty} E_n(f) \stackrel{u}{=} 0 \text{ in } [a, b]$$

□

7.1 Spline

Definition 7.10. Sia data una funzione $f : [a, b] \rightarrow \mathbb{R}$.

L'insieme delle spline di ordine k e' definito come:

$$S_k = \{s \in C^{k-1}([a, b]) \mid s|_{[x_{i-1}, x_i]} \in \mathbb{R}[x]_k, x_{i-1}, x_i \in \Delta_n \quad \forall i = 1, \dots, n \text{ con } \Delta_n \text{ partizione di } [a, b] \text{ di } n \text{ intervallini}\}$$

In pratica, $s \in S_k$ e' una funzione, con derivata $k-1$ continua, e in ogni intervallino $[x_{i-1}, x_i]$ e' un polinomio di grado $\leq k$.

Sia $s \in S_k$, definiamo

Spline naturale Se $s''(a) = s''(b) = 0$

Spline periodica Se

$$\begin{aligned} f(a) &= f(b) \\ s'_+(a) &= s'_-(b) \\ s''_+(a) &= s''_-(b) \end{aligned}$$

Di seguito esamineremo solo le spline naturali di S_3 .

Theorem 7.11. L'algoritmo che costruisce la spline naturale $s \in S_3$ e':

1. Si risolve il seguente sistema lineare $AM = B$, nell'incognita M , usando l'algoritmo Gauss

naive [5.2,pg.15] (nota⁸)

$$\begin{aligned}
A &\in \mathbb{R}^{n-1 \times n-1}, \quad M, B \in \mathbb{R}^{n-1 \times 1} \\
A_{ij} &= \begin{cases} 2(h_{i+1} + h_i) & i = j \\ h_{i+1} & j = i + 1 \\ h_i & j = i - 1 \\ 0 & \text{altrimenti} \end{cases} \\
B_i &= 6 \left(\frac{i+1-y_i}{h_{i+1}} + \frac{y_{i-1}-y_i}{h_i} \right) \quad \forall i = 1, \dots, n-1 \\
AM &= B_i
\end{aligned}$$

2. Si calcolano i seguenti valori:

$$\begin{aligned}
h_i &= x_i - x_{i-1} \quad \forall i = 1, \dots, n \\
\begin{cases} A_i = (M_{i-1} - M_i) \frac{h_i}{6} - \frac{-y_i + y_{i-1}}{h_i} \\ B_i = -\frac{1}{6} M_i h_i^2 + y_i \end{cases}
\end{aligned}$$

3. La funzione s risulta essere

$$\begin{aligned}
s(x) &= \begin{cases} s_1(x) & x \in [x_0, x_1] \\ s_2(x) & x \in [x_1, x_2] \\ \dots \\ s_n(x) & x \in [x_{n-1}, x_n] \end{cases} \\
s_i(x) &= M_i \frac{(x - x_{i-1})^3}{6h_i} + M_{i-1} \frac{(x_i - x)^3}{6h_i} + A_i(x - x_i) + B_i
\end{aligned}$$

Proof: Sia $s \in S_3$, fissiamo un Δ_n e poniamo

$$\begin{aligned}
M_k &= s''(x_k) \quad \forall k = 0, \dots, n \\
s_i &= s|_{[x_{i-1}, x_i]} \quad \forall i = 1, \dots, n \\
h_i &= x_i - x_{i-1} \quad \forall i = 1, \dots, n
\end{aligned}$$

Esaminiamo s_i

$$s_i \in \mathbb{R}[x]_3 \Rightarrow s_i = a_0 + a_1x + a_2x^2 + a_3x^3 \Rightarrow s_i'' = 2a_2 + 6a_3x$$

$$\begin{cases} s_i''(x_i) = M_i \\ s_i''(x_{i-1}) = M_{i-1} \\ \deg s_i'' = 1 \end{cases} \quad \underbrace{\Rightarrow}_{\text{retta passante per due punti}} \quad s_i''(x) = M_i \frac{x - x_{i-1}}{x_i - x_{i-1}} + M_{i-1} \frac{x_i - x}{x_i - x_{i-1}} = M_i \frac{x - x_{i-1}}{h_i} + M_{i-1} \frac{x_i - x}{h_i}$$

La funzione s'' risulta essere

$$s''(x) = \begin{cases} s_1''(x) & x \in [x_0, x_1] \\ s_2''(x) & x \in [x_1, x_2] \\ \dots \\ s_n''(x) & x \in [x_{n-1}, x_n] \end{cases}$$

⁸in realta' la matrice A e' tri-diagonale, e quindi esistono algoritmi adatti che risolvono il sistema in $O(5n)$

ed e' gia' continua, infatti, se esaminiamo il punto di raccordo x_i

$$\begin{aligned}\lim_{x \rightarrow x_i^-} s''(x) &= \lim_{x \rightarrow x_i^-} s''_i(x) \underbrace{=}_{s''_i(x) \text{ e' continua}} s''_i(x_i) = M_i \\ \lim_{x \rightarrow x_i^+} s''(x) &= \lim_{x \rightarrow x_{i+1}^-} s''_{i+1}(x) = s''_{i+1}(x_i) = M_i \\ \lim_{x \rightarrow x_i^-} s''(x) &= \lim_{x \rightarrow x_i^+} s''(x)\end{aligned}$$

Integrando $s''_i(x)$ ricaviamo $s'_i(x)$ e $s_i(x)$ (nota⁹):

$$\begin{aligned}s'_i(x) &= M_i \frac{(x - x_{i-1})^2}{2h_i} - M_{i-1} \frac{(x_i - x)^2}{2h_i} + A_i \\ s_i(x) &= M_i \frac{(x - x_{i-1})^3}{6h_i} + M_{i-1} \frac{(x_i - x)^3}{6h_i} + A_i x + B'_i = M_i \frac{(x - x_{i-1})^3}{6h_i} + M_{i-1} \frac{(x_i - x)^3}{6h_i} + A_i(x - x_i) + B_i\end{aligned}$$

Nota: Imponiamo il passaggio di s_i per i punti $(x_i, y_i = f(x_i))$, (x_{i-1}, y_{i-1}) :

$$\begin{cases} s_i(x_{i-1}) = y_{i-1} \\ s_i(x_i) = y_i \end{cases} \Leftrightarrow \begin{cases} A_i = (M_{i-1} - M_i) \frac{h_i}{6} - \frac{-y_i + y_{i-1}}{h_i} \\ B_i = -\frac{1}{6} M_i h_i^2 + y_i \end{cases}$$

Imponiamo la continuita' di $s'(x)$ nei punti di raccordo:

$$\begin{aligned}\lim_{x \rightarrow x_i^-} s'(x) &= \lim_{x \rightarrow x_i^+} s'(x) \Leftrightarrow s'_i(x_i) = s'_{i+1}(x_i) \Leftrightarrow \\ &\Leftrightarrow 2(h_{i+1} + h_i) M_i + h_{i+1} M_{i+1} + h_i M_{i-1} = 6 \left(\frac{i+1-y_i}{h_{i+1}} + \frac{y_{i-1}-y_i}{h_i} \right) \quad (1)\end{aligned}$$

Considerando che stiamo usando la condizione delle spline naturali

$$s''(a) = s''(b) = 0 \Leftrightarrow s''(x_0) = s''(x_n) = 0 \Leftrightarrow M_0 = M_n = 0$$

al variare di $i = 1, \dots, n-1$, nella (1) otteniamo un sistema lineare nelle $n-1$ incognite M_i e $n-1$ equazioni:

$$A \in \mathbb{R}^{(n-1) \times (n-1)}$$

$$\begin{aligned}A_{ij} &= \begin{cases} 2(h_{i+1} + h_i) & i = j \\ h_{i+1} & j = i+1 \\ h_i & j = i-1 \\ 0 & \text{altrimenti} \end{cases} \\ \bar{B}_i &= 6 \left(\frac{i+1-y_i}{h_{i+1}} + \frac{y_{i-1}-y_i}{h_i} \right) \quad \forall i = 1, \dots, n-1 \\ AM &= \bar{B}_i\end{aligned}$$

Poiche'

$$2(h_{i+1} + h_i) \underbrace{\geq}_{h_i \geq 0} h_{i+1} + h_i$$

si ha che A e' a predominanza diagonale stretta e quindi (vedi [5.3,pg.18]) possiamo usare l'algoritmo Gauss naive (vedi [5.2,pg.15]) per risolvere $AM = b$. $\square$

⁹ stiamo sfruttando la seguente proprieta':

$$\begin{aligned}\int (x-a)^k dx &\underbrace{=}_{y=x-a} \int y^k dy = \frac{y^{k+1}}{k+1} = \frac{(x-a)^{k+1}}{k+1} \\ \int (a-x)^k dx &\underbrace{=}_{y=a-x} \int -y^k dy = \frac{-y^{k+1}}{k+1} = \frac{-(a-x)^{k+1}}{k+1}\end{aligned}$$

Proposition 7.12. Sia $s \in S_3$ la spline naturale di $f \in C^4([a, b])$, si ha

$$|f(x) - s(x)| \leq \max_{\xi \in [x_0, x_n]} f^{(4)}(\xi) \frac{5}{384} (\max_j h_j)^4 \quad \forall x \in [x_0, x_n]$$

$$h_j = x_j - x_{j-1} \quad \forall j = 1, \dots, n$$

Questo assicura la buona approssimazione delle spline al diminuire di $\max_j h_j$ e quindi all'aumentare di n .

8 Metodo dei minimi quadrati

Proposition 8.1. 1. Siano $m, n \in \mathbb{N}$, $m < n$

2. Siano dati i punti $(x_i, y_i) \in \mathbb{R}^2$, $i = 1, \dots, n$ e i rispettivi vettori $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$

3. le funzioni $\varphi_j(x) : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ $j = 1, \dots, m$, con $x_i \in I \quad \forall i = 1, \dots, n$

4. Sia $\mathbf{a} = (a_1, \dots, a_m) \in \mathbb{R}^m$. Poniamo

$$\varphi(x, \mathbf{a}) = \sum_{j=1}^m a_j \varphi_j(x)$$

5. Sia

$$\phi \in \mathbb{R}^{n \times m}, \quad \phi_{jk} = \varphi_k(x_j)$$

$$S \in \mathbb{R}^{m \times m}, \quad S_{jk} = \sum_{i=1}^n \phi_{ij} \phi_{ik} = \sum_{i=1}^n \varphi_j(x_i) \varphi_k(x_i), \quad \Leftrightarrow S = {}^t \phi \phi$$

$$\mathbf{b} \in \mathbb{R}^m, \quad b_j = \sum_{i=1}^n ({}^t \phi)_{ji} y_i = \sum_{i=1}^n \varphi_j(x_i) y_i, \quad \Leftrightarrow \mathbf{b} = {}^t \phi \mathbf{y}$$

Con ϕ^k denotiamo la colonna k -esima di ϕ .

Supponendo che $\det S \neq 0$ il sistema

$$S\mathbf{a} = \mathbf{b} \Leftrightarrow {}^t \phi \phi \mathbf{a} = {}^t \phi \mathbf{y}$$

ammette una ed una sola soluzione nell'incognita $\mathbf{a}$. Questa soluzione minimizza la seguente funzione:

$$E(\mathbf{v}) = \|\varepsilon(\mathbf{v})\|_2^2 : \mathbb{R}^m \longrightarrow \mathbb{R}$$

Dove

$$\varepsilon(\mathbf{v}) : \mathbb{R}^m \longrightarrow \mathbb{R}^n$$

$$\varepsilon(\mathbf{v})_i = y_i - \varphi(x_i, \mathbf{v})$$

$E(\mathbf{v})$ indica di quanto la funzione $\varphi(x_i, \mathbf{v})$ stia sbagliando nell'approssimare i punti (x_i, y_i) . Infatti, se $\varepsilon(\mathbf{v}) = \mathbf{0}$, si avrebbe

$$y_i - \varphi(x_i, \mathbf{v}) = \varepsilon(\mathbf{v})_i = 0 \Leftrightarrow \varphi(x_i, \mathbf{v}) = y_i \quad \forall i = 1, \dots, n$$

e quindi $\varphi(x, \mathbf{v})$ sarebbe la funzione interpolante dei punti (x_i, y_i) . Poiche' $m < n$, in generale non esistera' un $\mathbf{v}$ tale che $\varepsilon(\mathbf{v}) = \underline{0}$. La soluzione $\mathbf{a}$ garantisce pero' che $E(\mathbf{a})$ sia il minimo errore, ovvero

$$E(\mathbf{a}) \leq E(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbb{R}^m$$

Proof: Sia $\mathbf{v} \in \mathbb{R}^m$.

$$\varepsilon(\mathbf{v})_i = y_i - \varphi(x_i, \mathbf{v}) = y_i - \sum_{j=1}^m v_j \varphi_j(x_i) = y_i - \sum_{j=1}^m v_j \phi_{ij} = y_i - \phi_i \mathbf{v}$$

dove ϕ_i e' la i -esima riga di ϕ e $\mathbf{v}$ e' inteso come colonna. Quindi

$$\varepsilon(\mathbf{v}) = \mathbf{y} - \phi \mathbf{v}$$

Dobbiamo minimizzare $\|\varepsilon(\mathbf{v})\|_2^2$. Osserviamo che

$$\|\varepsilon(\mathbf{v})\|_2^2 = \|\mathbf{y} - \phi \mathbf{v}\|_2^2 = d(\phi \mathbf{v}, \mathbf{y})^2$$

quindi dobbiamo trovare $\mathbf{v}$ in modo tale che la distanza tra i due vettori $\phi \mathbf{v}, \mathbf{y}$ sia minima. Poiche'

$$\phi \mathbf{v} = \sum_{i=1}^m \phi^i v_i$$

con ϕ^i la colonna i -esima di ϕ

possiamo interpretare $\phi \mathbf{v}$ come una combinazione lineare dei vettori ϕ^i , ovvero

$$\pi_n = \underbrace{\mathcal{L}(\phi^1, \dots, \phi^n)}_{\text{spazio generato}} = \{\phi \mathbf{v} \mid \mathbf{v} \in \mathbb{R}^m\}$$

Se $n = 2$, π_2 diventa un piano e il problema si puo' cosi' interpretare: trovare quel punto $\phi \mathbf{a}$ nel piano π_2 la cui distanza da $\mathbf{y}$ (che e' un punto dello spazio) sia minima. Intuitivamente, $\phi \mathbf{a}$ sembra essere la proiezione ortogonale di $\mathbf{y}$ su π_2 . Dimosteremo, nel caso generale, che e' proprio cosi'.

(1) Sia $\phi \mathbf{a}$ la proiezione di $\mathbf{y}$ su π_n . Si ha

$$\|\mathbf{y} - \phi \mathbf{a}\|_2^2 \leq \|\mathbf{y} - \phi \mathbf{v}\|_2^2 \quad \forall \mathbf{v} \in \mathbb{R}^m$$

ovvero

$$\|\varepsilon(\mathbf{a})\|_2^2 \leq \|\varepsilon(\mathbf{v})\|_2^2 \quad \forall \mathbf{v} \in \mathbb{R}^m$$

Infatti,

$\mathbf{y} - \phi \mathbf{a}$ e' il vettore che congiunge $\phi \mathbf{a}$ con $\mathbf{y}$ (*)

$\phi \mathbf{a} \in \pi_n$ e' la proiezione di $\mathbf{y}$ su π_n . Si ha $\underbrace{(\mathbf{y} - \phi \mathbf{a})}_{(*)} \perp \pi_n \Leftrightarrow \varepsilon(\mathbf{a}) \perp \pi_n \Leftrightarrow \varepsilon(\mathbf{a}) \perp \phi \mathbf{v} \quad \forall \mathbf{v} \in \mathbb{R}^m \Leftrightarrow$

$$\Leftrightarrow \underbrace{\varepsilon(\mathbf{a}) \cdot \phi \mathbf{v}}_{\text{prodotto scalare}} = 0 \quad \forall \mathbf{v} \in \mathbb{R}^m \quad (1)$$

$\mathbf{y} - \phi \mathbf{v} = \mathbf{y} - \phi \mathbf{a} + \phi \mathbf{a} - \phi \mathbf{v} = \mathbf{y} - \phi \mathbf{a} + \phi(\mathbf{a} - \mathbf{v}) = \varepsilon(\mathbf{a}) + \phi(\mathbf{a} - \mathbf{v})$ (*)

$$\|\mathbf{y} - \phi \mathbf{v}\|_2^2 = (\mathbf{y} - \phi \mathbf{a}) \cdot (\mathbf{y} - \phi \mathbf{a}) + \underbrace{2(\mathbf{y} - \phi \mathbf{a}) \cdot \phi(\mathbf{a} - \mathbf{v})}_{(*)} + \phi(\mathbf{a} - \mathbf{v}) \cdot \phi(\mathbf{a} - \mathbf{v}) =$$

$$= \varepsilon(\mathbf{a}) \cdot \varepsilon(\mathbf{a}) + 2\varepsilon(\mathbf{a}) \cdot \phi(\mathbf{a} - \mathbf{v}) + \phi(\mathbf{a} - \mathbf{v}) \cdot \phi(\mathbf{a} - \mathbf{v}) \stackrel{(1)}{=} \varepsilon(\mathbf{a}) \cdot \varepsilon(\mathbf{a}) + \phi(\mathbf{a} - \mathbf{v}) \cdot \phi(\mathbf{a} - \mathbf{v}) = \|\varepsilon(\mathbf{a})\|_2^2 + \|\phi(\mathbf{a} - \mathbf{v})\|_2^2$$

$$\|\mathbf{y} - \phi \mathbf{v}\|_2^2 = \|\varepsilon(\mathbf{a})\|_2^2 + \|\phi(\mathbf{a} - \mathbf{v})\|_2^2 \Leftrightarrow \|\varepsilon(\mathbf{a})\|_2^2 = \|\mathbf{y} - \phi \mathbf{v}\|_2^2 - \|\phi(\mathbf{a} - \mathbf{v})\|_2^2 \Rightarrow \|\varepsilon(\mathbf{a})\|_2^2 \leq \|\mathbf{y} - \phi \mathbf{v}\|_2^2 = \|\varepsilon(\mathbf{v})\|_2^2$$

Quindi per l'arbitrarieta' di $\mathbf{v}$ si ha la tesi.

(2) Determiniamo $\mathbf{a}$

Per la (1), $\varepsilon(\mathbf{a}) \perp \pi_n$

$$\varepsilon(\mathbf{a}) \perp \pi_n \Leftrightarrow \varepsilon(\mathbf{a}) \perp \phi^i \quad \forall i = 1, \dots, n \Leftrightarrow \phi^i \cdot \varepsilon(\mathbf{a}) = 0 \quad \forall i = 1, \dots, n \Leftrightarrow$$

[GII, 2.23, pg. 40]

$$\Leftrightarrow {}^t\phi \varepsilon(\mathbf{a}) = \underline{0} \Leftrightarrow {}^t\phi(\mathbf{y} - \phi \mathbf{a}) = \underline{0} \Leftrightarrow {}^t\phi \phi \mathbf{a} = {}^t\phi \mathbf{y} \quad (2)$$

$$({}^t\phi \phi)_{jk} = \sum_{i=1}^n ({}^t\phi)_{ji} \phi_{ik} = \sum_{i=1}^n \phi_{ij} \phi_{ik} = S_{jk} \Rightarrow {}^t\phi \phi = S$$

$$({}^t\phi \mathbf{y})_j = \sum_{i=1}^n ({}^t\phi)_{ji} y_i = \mathbf{b}_j \Rightarrow {}^t\phi \mathbf{y} = \mathbf{b}$$

$$(2) \Leftrightarrow S \mathbf{a} = \mathbf{b} \quad (2.1)$$

Quindi $\mathbf{a}$ e' soluzione del sistema (2.1). Poiche' per Hp $\det S \neq 0$, (2.1) ammette un'unica soluzione.

Nota: il sistema (2) viene chiamato *sistema normale*.

(3) Dimostrazione alternativa

Se le φ_i sono derivabili e se le ϕ^i sono vettori linearmente indipendenti si puo' procedere utilizzando il metodo di ricerca dei minimi relativi di una funzione reale:

1. Si impone $\nabla E(\mathbf{v}) = \underline{0}$. Con opportuni calcoli si ha

$$\nabla E(\mathbf{v}) = \underline{0} \Leftrightarrow S \mathbf{a} = \mathbf{b}$$

Poiche' $\det S \neq 0$ l'unico punto critico e' $\mathbf{a}$.

2. Si verifica che l'Hessiana $H(\mathbf{a})$ e' definita positiva, e quindi $\mathbf{a}$ e' un punto di minimo (l'unico). $\square$

Proposition 8.2. Supponiamo di essere nella situazione della prop [8.1, pg. 49]. Stavolta, per ogni punto (x_i, y_i) abbiamo a disposizione un numero $\omega_i > 0$ che rappresenta il peso del punto stesso. Vogliamo che la funzione $\varphi(x, \mathbf{a})$ sia piu' precisa nei punti che hanno maggior peso. Per far questo, basta definire un nuovo $\varepsilon(\mathbf{v})_i$:

$$\varepsilon_\omega(\mathbf{v})_i = \omega_i(y_i - \varphi(x_i, \mathbf{v}))$$

Ovviamente le quantita' in gioco cambieranno:

$$\phi \in \mathbb{R}^{n \times m}, \quad \phi_{jk} = \varphi_k(x_j)$$

$$\Omega = \text{diag}(\omega_1, \omega_2, \dots, \omega_n) \in \mathbb{R}^{n \times n}$$

$$S' \in \mathbb{R}^{m \times m}, \quad S'_{ij} = \sum_{k=1}^n \phi_{ki} \omega_k^2 \phi_{kj} = \sum_{k=1}^n \varphi_i(x_k) \varphi_j(x_k) \omega_k^2 \Leftrightarrow S' = {}^t\phi^t \Omega \phi$$

$$\mathbf{b}' \in \mathbb{R}^m, \quad b'_j = \sum_{k=1}^n \phi_{kj} \omega_k^2 y_k = \sum_{k=1}^n \varphi_j(x_k) \omega_k^2 y_k \Leftrightarrow \mathbf{b}' = {}^t\phi^t \Omega \mathbf{y}$$

Supponendo che $\det S' \neq 0$ il sistema

$$S' \mathbf{a} = \mathbf{b}' \Leftrightarrow {}^t\phi^t \Omega \phi \mathbf{a} = {}^t\phi^t \Omega \mathbf{y}$$

ammette una ed una sola soluzione nell'incognita $\mathbf{a}$. Questa soluzione minimizza la funzione:

$$E_\omega(\mathbf{v}) = \|\varepsilon_\omega(\mathbf{v})\|_2^2$$

Proof: Procedendo come nella dimostrazione della prop [8.1,pg.49]:

$$\varepsilon_\omega(\mathbf{v})_i = \omega_i(y_i - \phi_i \mathbf{v})$$

$$\varepsilon_\omega(\mathbf{v}) = \Omega(\mathbf{y} - \phi \mathbf{v})$$

dove

$$\Omega = \begin{bmatrix} \omega_1 & 0 & \dots & 0 \\ 0 & \omega_2 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & \omega_n \end{bmatrix}$$

$$\|\varepsilon_\omega(\mathbf{v})\|_2^2 = \|\Omega(\mathbf{y} - \phi \mathbf{v})\|_2^2 = (\Omega(\mathbf{y} - \phi \mathbf{v})) \cdot (\Omega(\mathbf{y} - \phi \mathbf{v})) = {}^t(\mathbf{y} - \phi \mathbf{v}) {}^t\Omega \Omega (\mathbf{y} - \phi \mathbf{v})$$

$${}^t\Omega \Omega = \begin{bmatrix} \omega_1^2 & 0 & \dots & 0 \\ 0 & \omega_2^2 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & \omega_n^2 \end{bmatrix}$$

Poiche' $\omega_i > 0$, ${}^t\Omega \Omega$ e' definita positiva; e' anche simmetrica. Quindi usando ${}^t\Omega \Omega$ possiamo definire il seguente prodotto scalare (vedi [GII,3,pg.56]) e la sua rispettiva norma:

$$x * y = {}^t\mathbf{x} {}^t\Omega \Omega \mathbf{y}$$

$$\|x\|_* = \sqrt{x * x}$$

Così' abbiamo:

$$\|\varepsilon_\omega(\mathbf{v})\|_2^2 = {}^t(\mathbf{y} - \phi \mathbf{v}) {}^t\Omega \Omega (\mathbf{y} - \phi \mathbf{v}) = (\mathbf{y} - \phi \mathbf{v}) * (\mathbf{y} - \phi \mathbf{v}) = \|\mathbf{y} - \phi \mathbf{v}\|_*^2 = d_*(\phi \mathbf{v}, \mathbf{y})^2$$

$$\|\varepsilon_\omega(\mathbf{v})\|_2^2 = \|\mathbf{y} - \phi \mathbf{v}\|_*^2 = \|\varepsilon(\mathbf{v})\|_*^2$$

Quindi, utilizzando il nuovo prodotto scalare, possiamo ripetere la dimostrazione della prop [8.1,pg.49]. Ripetiamo qui di seguito, in forma sintetica, tutti i passaggi.

(1) Sia $\phi \mathbf{a}$ proiezione di $\mathbf{y}$ su π_n . Si ha

$$\|\varepsilon(\mathbf{a})\|_*^2 \leq \|\varepsilon(\mathbf{v})\|_*^2 \quad \forall \mathbf{v} \in \mathbb{R}^m$$

Infatti,

$\mathbf{y} - \phi \mathbf{a}$ e' il vettore che congiunge $\phi \mathbf{a}$ con $\mathbf{y}$ (*)

$$\phi \mathbf{a} \in \pi_n \text{ e' la proiezione di } \mathbf{y} \underbrace{\Leftrightarrow (\mathbf{y} - \phi \mathbf{a}) \perp \pi_n}_{(*)} \Leftrightarrow \varepsilon(\mathbf{a}) \perp \pi_n \Leftrightarrow \varepsilon(\mathbf{a}) \perp \phi \mathbf{v} \quad \forall \mathbf{v} \in \mathbb{R}^m \Leftrightarrow$$

$$\Leftrightarrow \varepsilon(\mathbf{a}) * \phi \mathbf{v} = 0 \quad \forall \mathbf{v} \in \mathbb{R}^m \quad (1)$$

$$\mathbf{y} - \phi \mathbf{v} = \mathbf{y} - \phi \mathbf{a} + \phi \mathbf{a} - \phi \mathbf{v} = \mathbf{y} - \phi \mathbf{a} + \phi(\mathbf{a} - \mathbf{v}) = \varepsilon(\mathbf{a}) + \phi(\mathbf{a} - \mathbf{v}) \quad (*)$$

$$\|\mathbf{y} - \phi \mathbf{v}\|_*^2 = \|\varepsilon(\mathbf{a})\|_*^2 + \|\phi(\mathbf{a} - \mathbf{v})\|_*^2$$

$$\|\varepsilon(\mathbf{a})\|_*^2 = \|\mathbf{y} - \phi \mathbf{v}\|_*^2 - \|\phi(\mathbf{a} - \mathbf{v})\|_*^2 \Rightarrow \|\varepsilon(\mathbf{a})\|_*^2 \leq \|\mathbf{y} - \phi \mathbf{v}\|_*^2 = \|\varepsilon(\mathbf{v})\|_*^2$$

Quindi per l'arbitrarietà di $\mathbf{v}$ si ha la tesi.

(2) Determiniamo $\mathbf{a}$

Per la (1), $\varepsilon(\mathbf{a}) \perp \pi_n$

$$\varepsilon(\mathbf{a}) \perp \pi_n \Leftrightarrow \phi^i * \varepsilon(\mathbf{a}) = 0 \quad \forall i = 1, \dots, n \quad \underbrace{\Leftrightarrow}_{\text{def. di } *} {}^t(\phi^i) {}^t\Omega\Omega\varepsilon(\mathbf{a}) = 0 \quad \forall i \Leftrightarrow$$

$$\Leftrightarrow {}^t\phi {}^t\Omega\Omega\varepsilon(\mathbf{a}) = \underline{0} \Leftrightarrow {}^t\phi {}^t\Omega\Omega(\mathbf{y} - \phi\mathbf{a}) = \underline{0} \Leftrightarrow {}^t\phi {}^t\Omega\Omega\phi\mathbf{a} = {}^t\phi {}^t\Omega\Omega\mathbf{y} \quad (2)$$

Sia $D = {}^t\Omega\Omega$

$$({}^t\phi {}^t\Omega\Omega\phi)_{jk} = ({}^t\phi D\phi)_{jk} =$$

$$(D\phi)_{ij} = \sum_{k=1}^n D_{ik} \phi_{kj} = D_{ii} \phi_{ij} = \omega_i^2 \phi_{ij}$$

$$= \sum_{k=1}^n ({}^t\phi)_{ik} (D\phi)_{kj} = \sum_{k=1}^n \phi_{ki} \omega_k^2 \phi_{kj} = \sum_{k=1}^n \varphi_i(x_k) \varphi_j(x_k) \omega_k^2 = S'_{ij}$$

$$({}^t\phi {}^t\Omega\Omega\mathbf{y})_j = ({}^t\phi D\mathbf{y})_j = \sum_{k=1}^n ({}^t\phi D)_{jk} y_k =$$

$$({}^t\phi D)_{ij} = \sum_{k=1}^n ({}^t\phi)_{ik} D_{kj} = \sum_{k=1}^n \phi_{ki} D_{kj} = \phi_{ji} D_{jj} = \phi_{ji} \omega_j^2$$

$$= \sum_{k=1}^n \phi_{kj} \omega_k^2 y_k = \sum_{k=1}^n \varphi_j(x_k) \omega_k^2 y_k = b'_j$$

$$(2) \Leftrightarrow S' \mathbf{a} = \mathbf{b}' \quad (2.1)$$

Quindi $\mathbf{a}$ e' soluzione del sistema (2.1). Poiche' per Hp $\det S' \neq 0$, (2.1) ammette un'unica soluzione.

□

Proposition 8.3. *Supponiamo stavolta di volere approssimare un intero grafico di funzione $(x, f(x)) \quad \forall x \in I = [a, b]$. Si procede analogamente a quanto visto in [8.1, pg.49]:*

1. Sia

$$S \in \mathbb{R}^{m \times m}, \quad S_{jk} = \int_a^b \varphi_j(x) \varphi_k(x) dx$$

$$\mathbf{b} \in \mathbb{R}^m, \quad b_j = \int_a^b \varphi_j(x) f(x) dx$$

Supponendo che $\det S \neq 0$ il sistema

$$S\mathbf{a} = \mathbf{b}$$

ammette una ed una sola soluzione nell'incognita $\mathbf{a}$. Questa soluzione minimizza la seguente funzione:

$$E(\mathbf{v}) = \int_a^b \varepsilon(x, \mathbf{v})^2 dx : \mathbb{R}^m \longrightarrow \mathbb{R}$$

Dove

$$\varepsilon(x, \mathbf{v}) : I \times \mathbb{R}^m \longrightarrow \mathbb{R}$$

$$\varepsilon(x, \mathbf{v}) = f(x) - \varphi(x, \mathbf{v})$$

9 Equazioni non lineari

Proposition 9.1. *Metodo di bisezione.*

Sia $f \in C^0([a, b])$, con $f(a)f(b) < 0$, e posto

$$\begin{aligned} a_0 &= a, \quad b_0 = b \\ c_k &= \frac{a_k + b_k}{2} \\ (a_{k+1}, b_{k+1}) &= \begin{cases} (c_k, b_k) & f(c_k)f(a_k) > 0 \\ (a_k, c_k) & f(c_k)f(a_k) < 0 \\ \text{non definita} & f(c_k)f(a_k) = 0 \end{cases} \end{aligned}$$

Si ha che

$$\begin{aligned} \exists k : f(c_k) = 0 \quad \vee \quad \lim_{n \rightarrow +\infty} c_n = \alpha \in]a, b[\\ f(\alpha) = 0 \end{aligned}$$

Ovvero, o un qualche c_k e' uno zero di $f(x)$, oppure c_n converge a uno zero di f in $[a, b]$.
Nel secondo caso, si ha

$$|b_k - a_k| < 2\varepsilon \Rightarrow |c_{k+1} - \alpha| < \varepsilon$$

Dove $|c_{k+1} - \alpha|$ e' l'errore di approssimazione della radice α dato da c_{k+1} .
L'algoritmo che trova c_k o una approssimazione di α , con errore $< \varepsilon$, e' quindi il seguente:

While :

$$c := \frac{a + b}{2}$$

Se $|b - a| < 2\varepsilon \vee f(c) = 0$:

break

Se $f(c)f(a) > 0$:

$$a := c$$

Se $f(c)f(a) < 0$:

$$b := c$$

$$\alpha := c$$

Proof: Se $\exists k : f(c_k) = 0$ non abbiamo nulla da dimostrare. Supponiamo allora che

$$\forall k \quad f(c_k) \neq 0 \quad (0)$$

Poiche' $f(a)f(b) < 0$ possiamo supporre $f(a) > 0, f(b) < 0$. Abbiamo allora

(1) $f(a_k) > 0 \quad \forall k$

Per induzione.

Caso base:

$$f(a_0) = f(a) > 0$$

Supponiamo $f(a_k) > 0$. Abbiamo tre casi

1. Se $f(c_k)f(a_k) > 0$, allora $a_{k+1} = c_k$. Quindi

$$f(c_k)f(a_k) > 0 \quad \underbrace{\Rightarrow}_{f(a_k) > 0} f(c_k) > 0$$

$$f(a_{k+1}) = f(c_k) > 0$$

2. Se $f(c_k)f(a_k) < 0$, allora $a_{k+1} = a_k$ e quindi la tesi e' vera.
 3. Il caso $f(c_k)f(a_k) = 0$ non si puo' verificare perche' per la (0) si ha $f(c_k) \neq 0$.
 (2) $f(b_k) < 0 \quad \forall k$

Si procede analogamente a passo 1

Come abbiamo visto, il caso non definito della successione (a_k, b_k) non puo' verificarsi, quindi (a_k, b_k) diventa una successione ben definita:

$$(a_{k+1}, b_{k+1}) = \begin{cases} (c_k, b_k) & f(c_k)f(a_k) > 0 \\ (a_k, c_k) & f(c_k)f(a_k) < 0 \end{cases}$$

a cui possiamo applicare tutti i teoremi di analisi.

- (3) $a_k \leq b_k \quad \forall k$

Caso base: $a_0 \leq b_0$

Supponiamo $a_k \leq b_k$. Possono verificarsi due casi:

1. Se $(a_{k+1}, b_{k+1}) = (c_k, b_k)$, allora

$$a_{k+1} = c_k = \frac{a_k + b_k}{2} \leq b_{k+1} = b_k \Leftrightarrow a_k + b_k \leq 2b_k \Leftrightarrow a_k \leq b_k \text{ vera per Hp induttiva}$$

2. Se $(a_{k+1}, b_{k+1}) = (a_k, c_k)$, allora

$$a_{k+1} = a_k \leq b_{k+1} = c_k = \frac{a_k + b_k}{2} \Leftrightarrow 2a_k \leq a_k + b_k \Leftrightarrow a_k \leq b_k \text{ vera per Hp induttiva}$$

- (4) $a_{k+1} \geq a_k$

Se $a_{k+1} = a_k$ allora la tesi e' vera.

Se $a_{k+1} = c_k$, si ha

$$a_{k+1} = c_k = \frac{a_k + b_k}{2} \geq a_k \Leftrightarrow a_k + b_k \geq 2a_k \Leftrightarrow b_k \geq a_k \text{ vera per passo 3}$$

- (5) $b_{k+1} \leq b_k$

Se $b_{k+1} = b_k$ allora la tesi e' vera.

Se $b_{k+1} = c_k$, si ha

$$b_{k+1} = c_k = \frac{a_k + b_k}{2} \leq b_k \Leftrightarrow a_k + b_k \leq 2b_k \Leftrightarrow a_k \leq b_k \text{ vera per passo 3}$$

- (6) $b_k - a_k = \frac{b-a}{2^k}$

$$b_0 - a_0 = b - a = \frac{b-a}{2^0}$$

Supponiamo la tesi vera per k .

Possono verificarsi due casi:

1. Se $(a_{k+1}, b_{k+1}) = (c_k, b_k)$, allora

$$b_{k+1} - a_{k+1} = b_k - c_k = b_k - \frac{a_k + b_k}{2} = \frac{b_k - a_k}{2} \underset{\text{Hp indutt.}}{=} \frac{\frac{b-a}{2^k}}{2} = \frac{b-a}{2^{k+1}}$$

2. Se $(a_{k+1}, b_{k+1}) = (a_k, c_k)$, allora si procede analogamente a prima.

Abbiamo:

$$a_0 \underbrace{\leq}_{\text{passo 4}} a_k \underbrace{\leq}_{\text{passo 3}} b_k \underbrace{\leq}_{\text{passo 5}} b_0 \quad (1)$$

a_k e' quindi una successione non decrescente, limitata superiormente, quindi converge ad un α , che per la (1) e' in $[a, b]$.

Analogamente b_k converge a β . Abbiamo che $\alpha = \beta$, infatti,

$$\begin{cases} \lim_{n \rightarrow +\infty} b_n - a_n \underbrace{=}_{\text{passo 6}} \lim_{n \rightarrow +\infty} \frac{b-a}{2^n} = 0 \\ a_n \leq b_n \end{cases} \Rightarrow \lim_{n \rightarrow +\infty} b_n = \lim_{n \rightarrow +\infty} a_n \Leftrightarrow \alpha = \beta \quad (2)$$

$$\lim_{n \rightarrow +\infty} c_n = \lim_{n \rightarrow +\infty} \frac{a_n + b_n}{2} = \alpha \quad (2.1)$$

Infine, poichè f è continua in $[a, b]$ e $\alpha \in [a, b]$, si ha

$$\lim f(a_n) = f(\lim a_n) = f(\alpha) \quad (3.1)$$

$$\lim f(b_n) = f(\lim b_n) = f(\alpha) \quad (3.2)$$

Per il passo 1 e il passo 2 si ha

$$f(a_k) > 0 \quad \forall k \Rightarrow \lim f(a_n) \geq 0 \underbrace{\Leftrightarrow}_{(3.1)} f(\alpha) \geq 0$$

$$f(b_k) < 0 \quad \forall k \Rightarrow \lim f(b_n) \leq 0 \underbrace{\Leftrightarrow}_{(3.1)} f(\alpha) \leq 0$$

$$\Rightarrow f(\alpha) = 0$$

che è la tesi.

(7) Dimostriamo che

$$|b_k - a_k| < 2\varepsilon \Rightarrow |c_{k+1} - \alpha| < \varepsilon$$

$$a_k \text{ non decrescente} \Rightarrow \lim a_k = \sup a_k \Leftrightarrow \alpha = \sup a_k \Rightarrow \alpha \geq a_k \quad \forall k$$

$$b_k \text{ non crescente} \Rightarrow \lim b_k = \inf a_k \Leftrightarrow \alpha = \inf a_k \Rightarrow \alpha \leq b_k \quad \forall k$$

$$c_k \text{ è il punto medio di } [a_k, b_k] \Rightarrow c_k \in]a_k, b_k[$$

$$\alpha, c_k \in]a_k, b_k[\Rightarrow |c_k - \alpha| \leq |b_k - a_k| \underbrace{\leq}_{\text{per Hp}} 2\varepsilon \quad (+)$$

$$|b_k - a_k| \underbrace{=}_{\text{passo 6}} \frac{b-a}{2^k} \underbrace{\leq}_{(+)} 2\varepsilon \Leftrightarrow \frac{b-a}{2^{k+1}} < \varepsilon \quad (*)$$

$$|c_{k+1} - \alpha| \underbrace{\leq}_{(+)} |b_{k+1} - a_{k+1}| \underbrace{=}_{\text{passo 6}} \frac{b-a}{2^{k+1}} \underbrace{\leq}_{(*)} \varepsilon$$

□

Proposition 9.2. *Metodo di Newton.*

Sia

1. $f \in C^2([a, b])$
2. $f(a)f(b) < 0$
3. $f'(x) \neq 0 \quad \forall x \in [a, b]$
4. $f''(x) > 0 \quad \forall x \in [a, b]$ oppure $f''(x) < 0 \quad \forall x \in [a, b]$

Definita la successione

$$x_0 = \begin{cases} a & f(a)f''(a) > 0 \\ b & f(b)f''(b) > 0 \end{cases}$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

che risulta essere ben definita ($x_n \in [a, b] \quad \forall n$), si ha

$$\lim_{n \rightarrow +\infty} x_n = \alpha \in [a, b]$$

dove α e' uno zero di f , ovvero $f(\alpha) = 0$.

Geometricamente la successione x_n si puo' interpretare come segue:

1. si calcola la retta tangente in x_k
2. x_{k+1} e' il punto di intersezione di tale retta con l'asse $y = 0$

Proof: Supponiamo di essere nel caso $f'' > 0$, $f(b)f''(b) > 0$, $x_0 = b$. Analogamente si dimostrano gli altri casi.

(1) $f(x_k) > 0 \quad \forall k$

Poiche' $f''(x) > 0 \quad \forall x \in [a, b]$, si ha che f e' convessa in ogni punto (volge sempre la concavita' verso l'alto, vedi [AI,13.1,pg.47]). Quindi nel punto x_k abbiamo:

$$\exists \delta > 0 : f(x) > f(x_k) + f'(x_k)(x - x_k) \quad \forall x \in]x_k - \delta, x_k + \delta[\setminus \{x_k\}$$

Vale il realta' un risultato piu' forte, la funzione f e' convessa e si ha:

$$f(x) > f(x_k) + f'(x_k)(x - x_k) \quad \forall x \neq x_k \quad (*)$$

Allora

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \Leftrightarrow f'(x_k)(x_{k+1} - x_k) + f(x_k) = 0$$

$$f(x_{k+1}) \underbrace{>}_{(*)} f(x_k) + f'(x_k)(x_{k+1} - x_k) = 0$$

$$f(x_{k+1}) > 0$$

(2) f e' crescente e $f' > 0$

$$f'' > 0, f(b)f''(b) > 0 \Rightarrow f(b) > 0 \quad \underbrace{\Rightarrow}_{f(a)f(b) < 0} f(a) < 0$$

$$f' > 0 \vee f' < 0 \text{ infatti, se per assurdo } \exists x, x' : f'(x)f'(x') < 0 \quad \underbrace{\Rightarrow}_{f' \text{ continua, thm esistenza zeri}} \exists \xi : f'(\xi) = 0 \text{ assurdo}$$

$$f' > 0 \vee f' < 0 \Rightarrow f \text{ monotona}$$

$$\begin{cases} f(a) < 0, f(b) > 0 \\ f \text{ monotona} \end{cases} \Rightarrow f \text{ crescente} \Rightarrow f' > 0$$

(3) $\exists! \alpha \in [a, b] : f(\alpha) = 0$, inoltre, $a < \alpha < x_{k+1} < x_k < x_0 = b$

Poiche' $f(a)f(b) < 0$, per il thm degli zeri, $\exists \alpha \in [a, b] : f(\alpha) = 0$. Se per assurdo $\exists \alpha' \neq \alpha : f(\alpha') = 0$, allora

$$f(\alpha) = f(\alpha') = 0, f \text{ continua} \quad \underbrace{\Rightarrow}_{\text{thm di Rolle}} \exists \xi \in [\alpha, \alpha'] : f'(\xi) = 0 \text{ assurdo}$$

Quindi α e' unico.

$$\begin{cases} f(\alpha) = 0, f(x_k) \underbrace{>}_\text{passo 1} 0 \\ f \text{ crescente} \end{cases} \Rightarrow \alpha < x_k$$

Per i passi precedenti abbiamo $f(x_k), f'(x_k) > 0$, e quindi

$$\frac{f(x_k)}{f'(x_k)} > 0 \Rightarrow x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} < x_k$$

$$x_{k+1} < x_k \Rightarrow \{x_k\} \text{ e' decrescente}$$

$$\alpha < x_k < x_0 = b \Rightarrow \{x_k\} \text{ e' limitata}$$

$$\Rightarrow \exists \lim_{n \rightarrow +\infty} x_n = l$$

Prendendo il limite in $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$, per la continuita' di f , abbiamo

$$l = l - \frac{f(l)}{f'(l)} \Leftrightarrow f(l) = 0$$

Per l'unicita' di α , si ha $l = \alpha$.

□

Proposition 9.3. *Metodo delle secanti.*

Questo metodo e' analogo a quello di Newton, ma al posto di usare la retta tangente, si usa una retta secante (che sia abbastanza vicina alla tangente). La successione diventa:

$$x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}$$

Proposition 9.4. *Metodo misto.*

Il metodo di bisezione converge sempre ma e' molto lento.

Il metodo di Newton non sempre converge, ma e' molto rapido.

Si puo' pensare quindi di combinare i due metodi. Ad esempio,

1. *Si usa il metodo di bisezione di per restringere l'intervallo di ricerca*
2. *Si applica il metodo di Newton, usando un massimo m di iterazioni.*
3. *Se il massimo m e' stato raggiunto e lo zero non e' stato trovato, si ritorna al punto 1*

9.1 Ricerca di punti fissi

Proposition 9.5.

$$\begin{cases} \phi : [a, b] \longrightarrow [a, b] \\ \phi \in C^0([a, b]) \end{cases} \Rightarrow \exists \alpha \in [a, b] : \alpha = \phi(\alpha)$$

Proof: Sia $F(x) = x - \phi(x)$. Se $F(a) = 0$, allora $\alpha = a$. Altrimenti

$$\phi(a) \in [a, b] \Rightarrow \phi(a) \geq a \Leftrightarrow a - \phi(a) \leq 0 \underbrace{\Leftrightarrow}_{F(a) \neq 0} a - \phi(a) < 0 \Leftrightarrow F(a) < 0$$

Se $F(b) = 0$ allora $\alpha = b$, altrimenti

$$\phi(b) \in [a, b] \Rightarrow \phi(b) \leq b \Leftrightarrow b - \phi(b) \geq 0 \underbrace{\Leftrightarrow}_{F(b) \neq 0} b - \phi(b) > 0 \Leftrightarrow F(b) > 0$$

Allora

$$\begin{cases} F(a) < 0, F(b) > 0 \\ \phi(x) \text{ continua} \Rightarrow F(x) \text{ continua} \end{cases} \underbrace{\Rightarrow}_{\text{thm esistenza zeri}} \exists \alpha \in]a, b[: F(\alpha) = 0 \Leftrightarrow \phi(\alpha) = \alpha$$

□

Proposition 9.6. *Sia*

1. $\phi : [a, b] \longrightarrow [a, b]$
2. $\phi \in C^0([a, b])$
3. *Si verifica almeno una delle seguenti condizioni:*

(a) ϕ e' Lipschitziana di costante $0 < L < 1$, ovvero

$$|\phi(x) - \phi(y)| \leq L|x - y| \quad \forall x, y \in [a, b]$$

(b) oppure $\phi \in C^1([a, b])$ con

$$|\phi'(x)| \leq L < 1 \quad \forall x \in [a, b]$$

4. $x_0 \in [a, b], x_{k+1} = \phi(x_k)$

allora

$$\begin{aligned} \exists! \alpha \in [a, b] : \alpha &= \phi(\alpha) \\ \lim_{n \rightarrow +\infty} x_n &= \alpha \end{aligned}$$

Proof: Per la prop [9.5,pg.58], sappiamo che α esiste. Supponiamo per assurdo che esistano $\alpha_1 \neq \alpha_2$ t.c. $\phi(\alpha_i) = \alpha_i$, allora

CASE: Caso (a): ϕ Lipschitziana

$$|\alpha_1 - \alpha_2| = |\phi(\alpha_1) - \phi(\alpha_2)| \leq L|\alpha_1 - \alpha_2| \quad \underbrace{<}_{L < 1, \alpha_1 \neq \alpha_2} |\alpha_1 - \alpha_2|$$

$$|\alpha_1 - \alpha_2| < |\alpha_1 - \alpha_2|$$

Assurdo.

CASE: Caso (b): $|\phi'(x)| < 1 \quad \forall x \in [a, b]$

Poiche' $\alpha_1 \neq \alpha_2$, a meno di riordinare gli indici, possiamo supporre $\alpha_1 < \alpha_2$. Applichiamo il teorema di Lagrange:

$$\phi(x) \in C^1([a, b]) \quad \underbrace{\Rightarrow}_{\text{Lagrange}} \quad \exists d \in]\alpha_1, \alpha_2[: \quad \phi(\alpha_2) - \phi(\alpha_1) = \phi'(d)(\alpha_2 - \alpha_1) \Leftrightarrow$$

$$\frac{\phi(\alpha_2) - \phi(\alpha_1)}{\alpha_2 - \alpha_1} = \phi'(d) \Leftrightarrow \frac{\alpha_2 - \alpha_1}{\alpha_2 - \alpha_1} = \phi'(d) \Leftrightarrow \phi'(d) = 1$$

assurdo contro ipotesi.

(1) $|x_n - \alpha| \leq L^n |x_0 - \alpha|$

Per induzione.

Base:

$$|x_0 - \alpha| \leq L^0 |x_0 - \alpha|$$

Passo induttivo:

CASE: Caso (a): ϕ Lipschitziana

$$|x_{n+1} - \alpha| = |\phi(x_n) - \phi(\alpha)| \leq L|x_n - \alpha| \quad \underbrace{\leq}_{\text{Hp indutt.}} LL^n |x_0 - \alpha| = L^{n+1} |x_0 - \alpha|$$

CASE: Caso (b)

Se $x_n = \alpha$, allora

$$x_{n+1} = \phi(x_n) = \phi(\alpha) = \alpha$$

$$|x_{n+1} - \alpha| = 0 \leq L^{n+1}|x_0 - \alpha|$$

Supponiamo $\alpha < x_n$. Applichiamo Lagrange a $\phi(x)$ nell'intervallo $[\alpha, x_n]$, ottenendo:

$$\exists d \in]\alpha, x_n[: \phi(x_n) - \phi(\alpha) = \phi'(d)(x_n - \alpha)$$

$$|x_{n+1} - \alpha| = |\phi(x_n) - \phi(\alpha)| = |\phi'(d)||x_n - \alpha| \leq L|x_n - \alpha| \leq \underbrace{LL^n}_{\text{Hp indutt.}}|x_0 - \alpha| = L^{n+1}|x_0 - \alpha|$$

Poiche' $0 \leq L < 1$ si ha $L^n \rightarrow 0$, quindi prendendo il limite in

$$0 \leq |x_n - \alpha| \leq L^n|x_0 - \alpha|$$

abbiamo (per il thm del confronto) che:

$$\lim_{n \rightarrow +\infty} |x_n - \alpha| = 0 \Leftrightarrow \lim_{n \rightarrow +\infty} x_n = \alpha$$

□

Proposition 9.7. *Sia*

1. $\phi : [a, b] \longrightarrow \mathbb{R}$
2. $\phi \in C^1([a, b])$ con

$$|\phi'(x)| \geq L > 1 \quad \forall x \in [a, b]$$

3. $x_0 \in [a, b], x_{k+1} = \phi(x_k)$
4. $\alpha \in [a, b], \phi(\alpha) = \alpha$
5. $x_0 \neq \alpha$

*allora, la successione x_n **non** converge al punto fisso α . Anzi, la successione x_n non e' nemmeno ben definita:*

$$\exists k : x_k \notin [a, b]$$

Proof:

$$(1) |x_n - \alpha| \geq L^n|x_0 - \alpha| > 0$$

Per induzione.

Base:

$$|x_0 - \alpha| \geq L^0|x_0 - \alpha| = |x_0 - \alpha| \underbrace{>}_{{x_0 \neq \alpha}} 0$$

Passo induttivo: Per Hp induttiva $|x_n - \alpha| > 0$ e quindi $x_n \neq \alpha$. Supponiamo $\alpha < x_n$.

Applichiamo Lagrange a $\phi(x)$ nell'intervallo $[\alpha, x_n]$, ottenendo:

$$\exists d \in]\alpha, x_n[: \phi(x_n) - \phi(\alpha) = \phi'(d)(x_n - \alpha)$$

$$|x_{n+1} - \alpha| = |\phi(x_n) - \phi(\alpha)| = |\phi'(d)||x_n - \alpha| \underbrace{\geq}_{|\phi'(x)| \geq L} L|x_n - \alpha| \underbrace{\geq}_{\text{Hp indutt.}} LL^n|x_0 - \alpha| = L^{n+1}|x_0 - \alpha| > 0$$

Quindi:

$$|x_n - \alpha| \geq L^n|x_0 - \alpha| \quad \forall n$$

e poiche' $L > 1$, prendendo il limite, si ha

$$\lim_{n \rightarrow +\infty} |x_n - \alpha| = +\infty \Rightarrow \lim_{n \rightarrow +\infty} x_n \neq \alpha$$

Infine, se per assurdo $x_n \in [a, b] \quad \forall n$, allora

$$x_n, \alpha \in [a, b] \quad \forall n \Rightarrow |x_n - \alpha| \leq |b - a| \quad \forall n \Rightarrow \lim_{n \rightarrow +\infty} |x_n - \alpha| \leq |b - a|$$

assurdo contro $\lim_{n \rightarrow +\infty} |x_n - \alpha| = +\infty$

□

Proposition 9.8. *Sia*

1. $\phi : [a, b] \longrightarrow \mathbb{R}$
2. $\phi \in C^1([a, b])$
3. $x_0 \in [a, b], x_{k+1} = \phi(x_k)$
4. $\alpha = \phi(\alpha)$

allora

$$|\phi'(\alpha)| < 1 \Rightarrow \exists \delta > 0 : (x_0 \in I_\delta \Rightarrow \lim_{n \rightarrow +\infty} x_n = \alpha)$$

$$\text{dove } I_\delta = [\alpha - \delta, \alpha + \delta] \subseteq [a, b]$$

Proof:

$$|\phi'(\alpha)| < 1 \Rightarrow \exists \eta : |\phi'(\alpha)| < \eta < 1$$

$$\left\{ \begin{array}{l} \phi' \text{ e' continua} \\ |\phi'(\alpha)| < \eta \end{array} \right. \xRightarrow{\text{permanenza del segno}} \exists \delta > 0 : |\phi'(x)| < \eta < 1 \quad \forall x \in I_\delta \quad (1)$$

Sia $x_0 \in I_\delta$. Se $\exists m : x_m = \alpha$, allora non abbiamo nulla da dimostrare dato che x_n diventa la successione costante $\{\alpha\}$. Supponiamo quindi che

$$\forall n : x_n \neq \alpha$$

(1) Dimostriamo che

$$x_n \in I_\delta, |x_n - \alpha| \leq |x_0 - \alpha| \quad \forall n$$

$$|x_n - \alpha| < \eta^n |x_0 - \alpha|$$

Per induzione.

Base:

$$|x_0 - \alpha| \leq \eta^0 |x_0 - \alpha|$$

Passo induttivo:

$x_n \in I_\delta \subseteq [a, b]$ per Hp induttiva. Supponiamo $x_n \neq b$ e $\alpha < x_n$

$$\phi \in C^1([a, b]) \xRightarrow{x_n \in [a, b], \alpha \in [a, b], \alpha \neq x_n} \phi \in C^1([\alpha, x_n]) \xRightarrow{\text{thm di Lagrange}} \exists \xi_n \in]\alpha, x_n[: \phi(x_n) - \phi(\alpha) = \phi'(\xi_n)(x_n - \alpha) \quad (*)$$

$$\xi_n \in]\alpha, x_n[\Leftrightarrow \alpha < \xi_n < x_n \leq \alpha + \delta$$

$$\alpha - \delta < \alpha < \xi_n < \alpha + \delta \Rightarrow \xi_n \in]\alpha - \delta, \alpha + \delta[= I_\delta$$

$$\xRightarrow{(1)} |\phi'(\xi_n)| < \eta < 1 \quad (**)$$

$$|x_{n+1} - \alpha| = |\phi(x_n) - \phi(\alpha)| \underset{(*)}{=} |\phi'(\xi_n)| |x_n - \alpha| \underset{x_n \neq \alpha, (**)}{<} \eta |x_n - \alpha| \leq$$

$$\underset{\text{Hp induttiva}}{\leq} \eta \eta^n |x_0 - \alpha| = \eta^{n+1} |x_0 - \alpha| \underset{\eta < 1}{<} |x_0 - \alpha| \underset{x_0 \in I_\delta}{\leq} \delta$$

$$|x_{n+1} - \alpha| \leq \delta \Rightarrow x_{n+1} \in I_\delta$$

$$|x_{n+1} - \alpha| < \eta^{n+1} |x_0 - \alpha|$$

Prendendo il limite in

$$0 \leq |x_n - \alpha| \leq \eta^n |x_0 - \alpha|$$

poiche' $0 \leq \eta < 1, \eta^n \rightarrow 0$, quindi per il thm del confronto

$$\lim_{n \rightarrow +\infty} |x_n - \alpha| = 0 \Leftrightarrow \lim_{n \rightarrow +\infty} x_n = \alpha$$

□

Proposition 9.9. *Supponiamo di volere risolvere l'equazione*

$$f(x) = g(x)$$

Se f ammette l'inversa, possiamo ricondurci alla ricerca di un punto fisso:

$$x = f^{-1}g(x) = \phi_1(x)$$

Se g ammette l'inversa, possiamo ricondurci a:

$$x = g^{-1}f(x) = \phi_2(x)$$

Scelta ϕ_1 o ϕ_2 possiamo cercare di applicare i thm fin'ora visti.

Supponendo che f, g ammettano le inverse, possiamo decidere quale delle due ϕ_i preferire, infatti

$$\begin{aligned} |\phi'_1(x)| < 1 \quad \forall x \in [a, b] &\Leftrightarrow |\phi'_2(x)| > 1 \quad \forall x \in [a, b] \\ |\phi'_1(x)| > 1 \quad \forall x \in [a, b] &\Leftrightarrow |\phi'_2(x)| < 1 \quad \forall x \in [a, b] \end{aligned}$$

Poiche' vogliamo sfruttare il thm [9.6,pg.59], sceglieremo quindi la ϕ_i t.c. $|\phi'_i(x)| < 1$. Infine, osserviamo che dal problema della ricerca di un punto fisso possiamo passare al problema della ricerca di uno zero di funzione:

$x = hg(x)$ ricerca del punto fisso

$h^{-1}(x) = g(x) \Leftrightarrow h^{-1}(x) - g(x) = 0$ ricerca di zero di funzione

Proof: Basta osservare che

$$\phi_1(x) = \phi_2^{-1}(x)$$

e applicare il thm di derivazione della funzioni inverse (vedi [AI,11.3,pg.41]):

$$\begin{aligned} \phi'_1(x) &= D[\phi_2^{-1}(x)] = \frac{1}{\phi'_2(\phi_2^{-1}(x))} \\ |\phi'_1(x)| &= \frac{1}{|\phi'_2(\phi_2^{-1}(x))|} \end{aligned}$$

□

9.2 Ordine di convergenza

Definition 9.10. Consideriamo una successione x_n che converge ad α , e sia $e_n = x_n - \alpha$. Per misurare quanto rapidamente la successione converge, poniamo

$$E_{n,p} = \frac{|e_{n+1}|}{|e_n|^p}$$

e diamo le seguenti definizioni

$$x_n \text{ ha ordine di convergenza (O.C.) } 1 \stackrel{\text{def}}{\Leftrightarrow} \lim_{n \rightarrow +\infty} E_{n,1} < 1$$

$$x_n \text{ ha O.C. } p \stackrel{\text{def}}{\Leftrightarrow} \lim_{n \rightarrow +\infty} E_{n,p} = l \neq 0, \quad l \in \mathbb{R}, \quad p \geq 1$$

$$x_n \text{ e' superlineare} \stackrel{\text{def}}{\Leftrightarrow} \lim_{n \rightarrow +\infty} E_{n,1} = 0$$

Proposition 9.11. *Sia*

1. $\phi : [a, b] \longrightarrow \mathbb{R}$
2. $x_0 \in [a, b], x_{k+1} = \phi(x_k) \in [a, b]$
3. $x_k \rightarrow \alpha \in [a, b], \alpha = \phi(\alpha)$
4. ϕ derivabile n volte, e sia

$$k = \min \left\{ k \in \mathbb{N} \mid \phi^{(k)}(\alpha) \neq 0 \right\}$$

con $1 < k \leq n$

5. $\phi^{(k)}(x)$ continua in α

Si ha

x_n ha ordine di convergenza k

Inoltre, se $k = 1$,

$$|\phi'(\alpha)| < 1, k = 1 \Rightarrow x_n \text{ ha ordine di convergenza } 1$$

Proof: Per la formula di Taylor con il resto di Lagrange (vedi [AII,7,pg.59]) si ha:

dato $x_n \in [a, b] \setminus \{\alpha\} \exists \xi_n \in]\min\{x_n, \alpha\}, \max\{x_n, \alpha\}[$ tale che

$$\phi(x_n) = \phi(\alpha) + \frac{\phi'(\alpha)}{1!}(x_n - \alpha) + \dots + \frac{D^{k-1}[\phi(\alpha)]}{(k-1)!}(x_n - \alpha)^{k-1} + \frac{D^k[\phi(\xi_n)]}{k!}(x_n - \alpha)^k$$

ma per Hp tutte le derivate di ordine $< k$ sono nulle in α , quindi

$$\begin{aligned} \phi(x_n) &= \phi(\alpha) + \frac{D^k[\phi(\xi_n)]}{k!}(x_n - \alpha)^k \Leftrightarrow \\ &\Leftrightarrow \frac{\phi(x_n) - \phi(\alpha)}{(x_n - \alpha)^k} = \frac{D^k[\phi(\xi_n)]}{k!} \Leftrightarrow \\ &\Leftrightarrow \frac{x_{n+1} - \alpha}{(x_n - \alpha)^k} = \frac{D^k[\phi(\xi_n)]}{k!} \Rightarrow \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^k} = \frac{|\phi^{(k)}(\xi_n)|}{k!} \end{aligned}$$

Prendendo i limiti abbiamo

$$\begin{aligned} &\begin{cases} x_n \rightarrow \alpha \\ \xi_n \in]\min\{x_n, \alpha\}, \max\{x_n, \alpha\}[\end{cases} \quad \xi_n \rightarrow \alpha \\ &\lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^k} = \lim_{n \rightarrow +\infty} \frac{|\phi^{(k)}(\xi_n)|}{k!} \stackrel{\phi^{(k)} \text{ continua in } \xi_n, \text{ [AII,8.31,pg.87]}}{=} \frac{|\phi^{(k)}(\alpha)|}{k!} \neq 0 \end{aligned}$$

Quindi, se $k > 1$ la successione ha O.C. p .

Supponiamo che $k = 1$, con $|\phi'(\alpha)| < 1$ allora

$$0 < |\phi'(\alpha)| < 1 \quad (1)$$

$$\lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^k} = |\phi'(\alpha)| < 1$$

e quindi x_n ha ordine di convergenza 1. □

Proposition 9.12. Poniamoci nelle ipotesi della prop [9.2,pg.56], e supponiamo che $\exists f'''(\alpha)$. Allora, l'ordine di convergenza della successione x_n del metodo di Newton e' 2.

Proof: Adoperiamo la prop [9.11,pg.62]. Nel caso del metodo di Newton ϕ e' la funzione

$$\phi(x) = x - \frac{f(x)}{f'(x)}$$

Ricordiamo che nelle Hp della prop [9.2,pg.56], $\phi''(x), \phi'(x)$ non sono mai nulle. Calcoliamo le derivate di ϕ :

$$\begin{aligned}\phi'(x) &= f(x) \frac{f''(x)}{f'(x)^2} \\ \phi'(\alpha) &\underbrace{=}_{f(\alpha)=0} 0 \\ \phi''(x) &= \frac{f''(x)}{f'(x)} + f(x) \left(-2 \frac{(f''(x))^2}{(f'(x))^3} + \frac{f'''(x)}{(f'(x))^2} \right) \\ \phi''(\alpha) &= \frac{f''(\alpha)}{f'(\alpha)} \neq 0\end{aligned}$$

□

Proposition 9.13. *Nel caso migliore, il metodo delle secanti (vedi [9.3,pg.58]) ha ordine $p = \frac{1+\sqrt{5}}{2}$.*

$$\begin{aligned}p^2 - p - 1 &= 0 \\ 1 < p < 2\end{aligned}$$

9.2.1 Criteri di arresto

Definition 9.14. In generale, un criterio di arresto e' una condizione P tale che

$$P \text{ vera} \Rightarrow \text{errore} < \varepsilon$$

dove ε e' la precisione di macchina.

Proposition 9.15. *Sia $x_{n+1} = \varphi(x_n)$, allora come criterio di arresto possiamo utilizzare la seguente condizione:*

$$|x_{n+1} - x_n| < \varepsilon$$

Non sempre pero' tale criterio e' soddisfacente, infatti, anche se $x_n \rightarrow \alpha$, con $\varphi(\alpha) = \alpha$, potrebbe accadere che $|x_{n+1} - x_n| \rightarrow \infty$ oppure che $|e_n| = |x_n - \alpha| > \varepsilon$.

Se la successione x_n viene usata per la ricerca di uno zero di $f(x)$ (come nel caso del metodo di Newton), allora un altro criterio e'

$$|f(x_n)| < \varepsilon$$

Anche questo criterio puo' non funzionare a dovere, infatti, anche se $|f(x_n)| < \varepsilon$ potrebbe accadere che $|e_n| = |x_n - \alpha| > \varepsilon$, oppure $|f(x_n)| > \varepsilon$ e $|e_n| < \varepsilon$. Nel secondo caso il metodo non si arresta, anche se potrebbe, e quindi compira' calcoli superflui.

Proposition 9.16. *Metodo di Aitken.*

Consideriamo la successione $x_n = \varphi(x_{n-1})$ che converge linearmente al punto fisso $\alpha = \varphi(\alpha)$. Supponiamo che $\varphi'(\alpha) \neq 0$, $x_n \neq \alpha \forall n$, allora vale il seguente criterio di arresto:

$$\left| x_n - \frac{x_{n+1}x_{n-1} - x_n^2}{x_{n+1} + x_{n-1} - 2x_n} \right| < \varepsilon$$

se la condizione e' verificata, il metodo ritornera' x_{n+1} .

Proof: Per la formula di Taylor con il resto di Lagrange (vedi [AII,7,pg.59]) si ha:
dato $x_n \in [a, b] \setminus \{\alpha\} \exists \xi_n \in]\min\{x_n, \alpha\}, \max\{x_n, \alpha\}[$ tale che

$$\begin{aligned}\varphi(x_n) &= \varphi(\alpha) + \frac{\varphi'(\alpha)}{1!}(x_n - \alpha) + \frac{D^2[\varphi(\xi_n)]}{2!}(x_n - \alpha)^2 \Leftrightarrow \\ \Leftrightarrow x_{n+1} - \alpha &= \frac{\varphi'(\alpha)}{1!}(x_n - \alpha) + \frac{D^2[\varphi(\xi_n)]}{2!}(x_n - \alpha)^2 \quad (1)\end{aligned}$$

Allo stesso modo:

$$x_n - \alpha = \frac{\varphi'(\alpha)}{1!}(x_{n-1} - \alpha) + \frac{D^2[\varphi(\xi_{n-1})]}{2!}(x_{n-1} - \alpha)^2 \quad (2)$$

Poiche' $x_n - \alpha \rightarrow 0$, per n sufficientemente grande, possiamo trascurare il resto nella formula di Taylor, ottenendo il sistema:

$$\begin{cases} x_{n+1} - \alpha \approx \varphi'(\alpha)(x_n - \alpha) \\ x_n - \alpha \approx \varphi'(\alpha)(x_{n-1} - \alpha) \end{cases} \quad (3)$$

Per Hp abbiamo che $\varphi'(\alpha) \neq 0, x_n \neq \alpha \forall n$, quindi

$$(3) \quad \underbrace{\Rightarrow}_{\text{dividendo membro a membro}} \frac{x_{n+1} - \alpha}{x_n - \alpha} \approx \frac{x_n - \alpha}{x_{n-1} - \alpha} \Leftrightarrow (x_{n+1} - \alpha)(x_{n-1} - \alpha) \approx (x_n - \alpha)(x_n - \alpha) \Rightarrow$$

$$\Rightarrow \alpha \approx \frac{x_{n+1}x_{n-1} - x_n^2}{x_{n+1} + x_{n-1} - 2x_n} = \alpha'$$

abbiamo cosi' ottenuto una approssimazione α' di α , migliore di x_n .
...DA CONTINUARE...

□

9.3 Sistemi non lineari

Proposition 9.17. Sia $X \subseteq \mathbb{R}^m$ e sia

$$\begin{aligned}f &: X \longrightarrow \mathbb{R}^m \\ f(x) &= (f_1(x), \dots, f_m(x)) \\ f_i &: \mathbb{R}^m \longrightarrow \mathbb{R}\end{aligned}$$

Risolvere $f(x) = \underline{0}$ equivale a risolvere il sistema

$$\begin{cases} f_1(x_1, \dots, x_m) = 0 \\ f_2(x_1, \dots, x_m) = 0 \\ \dots \\ f_m(x_1, \dots, x_m) = 0 \end{cases}$$

Supponendo $f \in \mathbb{C}^2(X)$ la generalizzazione del metodo di Newton diventa:

$n := 0$

While :

$$J(x^{(n)})\Delta x = -f(x^{(n)}) \quad \text{ sistema lineare nell'incognita } \Delta x$$

$$x^{(n+1)} = \Delta x + x^{(n)}$$

$n := n + 1$

Se $\|\Delta x\| < \varepsilon$

la soluzione approssimata con errore $< \varepsilon$ e' $x^{(n+1)}$

break

dove

$J(x)$ e' la matrice Jacobiana di f calcolata in x :

$$J_{ij}(x) = \left(\frac{\partial f_i}{\partial x_j} \right)_x \quad \forall i, j = 1, \dots, m$$

Proof: Enunciamo solo la struttura della dimostrazione.

Sia $\alpha \in X$: $f(\alpha) = 0$. Utilizziamo la formula di Taylor generalizzata:

$$f(\alpha) = f(x^{(n)}) + f'(x^{(n)})(\alpha - x^{(n)}) + \frac{1}{2} f''(\xi^{(n)})(\alpha - x^{(n)})^2$$

dove

$$f'(x^{(n)}) = J(x^{(n)})$$

$$f''_{ijk}(\xi^{(n)}) = \left(\frac{\partial^2 f_i}{\partial x_j \partial x_k} \right)_{\xi^{(n)}}$$

(nota¹⁰) e trascuriamo il resto poiche' e' un termine quadratico:

$$f(\alpha) \approx f(x^{(n)}) + f'(x^{(n)})(\alpha - x^{(n)})$$

Poiche' $f'(x^{(n)}) = J(x^{(n)})$ e $f(\alpha) = 0$ abbiamo:

$$0 \approx f(x^{(n)}) + f'(x^{(n)})(\alpha - x^{(n)})$$

Ponendo $\Delta x = \alpha - x^{(n)}$, abbiamo cosi' un sistema lineare. La cui soluzione dara' la seguente approssimazione:

$$\alpha - x^{(n)} \approx \Delta x$$

$$\alpha \approx \Delta x + x^{(n)}$$

Quindi poniamo $x^{(n+1)} = \Delta x + x^{(n)}$ e reiteriamo. □

9.4 Massimi e minimi

Per trovare il minimo di una funzione f con derivata continua possiamo usare i metodi forniti dall'analisi, che consistono principalmente nel trovare gli zeri di f' . Per trovare gli zeri possiamo usare i metodi visti in [9,pg.54].

In questo paragrafo vedremo dei metodi piu' generali che si possono applicare anche a funzioni non derivabili.

Proposition 9.18. *Metodo della sezione aurea.*

Sia $f : [a, b] \longrightarrow \mathbb{R}$ (nota¹¹) e unimodale, ovvero,

$$\exists x_* \in [a, b] : \begin{cases} f \text{ e' strettamente decrescente in } [a, x_*] \\ f \text{ e' strettamente crescente in } [x_*, b] \end{cases}$$

¹⁰ $f''_{ijk} \in \mathbb{R}^{m \times m \times m}$ e' una matrice a tre dimensioni. In generale, se $v \in \mathbb{R}^{m_3 \times 1}$ e' un vettore colonna e $A \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ e' una matrice a tre dimensioni, A_i sara' la i -esima matrice $m_2 \times m_3$ di A , Av sara' una matrice $m_1 \times m_2$, la cui riga i -esima e' $A_i v$.

¹¹ non stiamo richiedendo che f sia continua

allora il seguente algoritmo restituisce il minimo di f con un errore inferiore a ε :

While

$$x_1 := a + (1 - r)(b - a)$$

$$x_2 := a + r(b - a)$$

Se $f(x_1) > f(x_2)$:

$$a := x_1$$

Altrimenti:

$$b := x_2$$

Se $b - a < \varepsilon$

$$x_* := a$$

break

dove $r \in]\frac{1}{2}, 1[$.

Come r possiamo scegliere la sezione aurea, ovvero la soluzione positiva dell'equazione:

$$p^2 + p - 1 = 0$$

$$p = \frac{-1 + \sqrt{5}}{2}$$

Così facendo, nella $k + 1$ -esima iterazione potremmo usare un x_1 o x_2 precedente e calcolare solo un nuovo x'_1 o x'_2 , ad esempio: $x'_1 = x_1$, $x_2 = a + r(b - a)$

Proof: Supponiamo che $x_* \neq a$, $x_* \neq b$, allora presi due punti distinti $x_1, x_2 \in [a, b]$, con $x_1 < x_2$, abbiamo

(1) $f(x_1) < f(x_2) \Rightarrow x_* \in [a, x_2]$

Se per assurdo $x_* \in]x_2, b]$, allora

$$x_1 < x_2 < x_* \quad \underbrace{\Rightarrow}_{f \text{ decrescente in } [a, x_*]} \quad \Rightarrow f(x_1) > f(x_2) \text{ assurdo}$$

Analogamente

(2) $f(x_1) > f(x_2) \Rightarrow x_* \in [x_1, b]$

Se per assurdo $x_* \in [a, x_1[$, allora

$$x_* < x_1 < x_2 \quad \underbrace{\Rightarrow}_{f \text{ crescente in } [x_*, b]} \quad \Rightarrow f(x_1) < f(x_2) \text{ assurdo}$$

Quindi, prendendo due punti in $[a, b]$ sappiamo decidere in quale sottointervallo si trova il minimo. A questo punto, basterà prendere due nuovi punti in questo sottointervallo e ripetere la procedura.

Sceghieremo in modo opportuno i punti x_1, x_2 :

$$x_1 = a + (1-r)(b-a) = b - r(b-a), \quad x_2 = a + r(b-a)$$

$$\text{con } r \in]\frac{1}{2}, 1[$$

$$a < x_1 < \frac{a+b}{2} < x_2 < b, \text{ infatti}$$

$$0 < (1-r)(b-a) < \frac{b-a}{2} < r(b-a) < b-a$$

$$0 < 1-r < \frac{1}{2} < r < 1$$

$$\begin{cases} 0 < 1-r < \frac{1}{2} \\ \frac{1}{2} < r < 1 \end{cases} \Leftrightarrow \begin{cases} 0 < \frac{1}{2} < r \\ \frac{1}{2} < r < 1 \end{cases} \quad \text{vera}$$

I due possibili sottointervalli sono:

$$[a, x_2], [x_1, b]$$

ed hanno la stessa ampiezza:

$$x_2 - a = b - x_1 \Leftrightarrow r(b-a) = b-a - (1-r)(b-a) = r(b-a) \Leftrightarrow 0=0$$

Come r possiamo scegliere la sezione aurea: $r = \frac{-1+\sqrt{5}}{2}$.

Sia $[a_k, b_k]$ l'intervallo della k -esima iterazione,

Per induzione, si ha che l'ampiezza dell'intervallo e': $r^k(b-a)$ (1)

$$x_* \in [a_k, b_k] \Rightarrow |e_k| = |x_* - a_k| \leq |b_k - a_k| \underbrace{=}_{(1)} r^k(b-a) \quad (1.1)$$

e poiche' $0 < r < 1$, per $k \rightarrow \infty$ la (1) converge a 0, e quindi $a_k \rightarrow x_*$.

Dalla (1.1) possiamo ricavare anche il seguente criterio di arresto:

$$|b_k - a_k| < \varepsilon$$

Supponiamo che il nuovo intervallo sia $[a, x_2]$, cioe' $a' = a$, $b' = x_2$. Se vogliamo risparmiare il calcolo di un punto nella nuova suddivisione, imponiamo che

$$x'_2 = x_1 \Leftrightarrow a' + r(b' - a') = a + (1-r)(b-a) \Leftrightarrow r(x_2 - a) = (1-r)(b-a) \Leftrightarrow$$

$$\Leftrightarrow r(r(b-a)) = (1-r)(b-a) \Leftrightarrow r^2 = 1-r$$

sostituendo x_2

□

Proposition 9.19. *Metodo della parabola.*

Si interpola la funzione f in tre punti, ottenendo cosi una parabola. Il minimo della parabola sara' una approssimazione del minimo di f .

Questo metodo puo' essere usato in congiunzione al metodo della sezione aurea.

10 Integrali

Definition 10.1. Data la funzione peso $w : [a, b] \rightarrow \mathbb{R}$, con $w(x) > 0 \quad \forall x \in [a, b] \setminus \{z_1, \dots, z_s\}$, $w(z_i) = 0 \quad \forall i = 1, \dots, s$, definiamo l'integrale pesato:

$$I : C^0([a, b]) \rightarrow \mathbb{R}$$

$$I(f) = \int_a^b w(x)f(x) dx$$

Scelto $n + 1$ nodi $x_i \in [a, b] \quad \forall i = 0, \dots, n$, con $x_0 < x_1 < \dots < x_n$, e scelto $n + 1$ pesi $w_i \in \mathbb{R} \quad \forall i = 0, \dots, n$, chiameremo formula di quadratura la seguente funzione:

$$I_n : C^0([a, b]) \longrightarrow \mathbb{R}$$

$$I_n(f) = \sum_{i=0}^n w_i f(x_i)$$

Proposition 10.2. *Dalla proprietà dell'integrale definito segue che I è una funzione lineare, ovvero prese $f, g \in C^0([a, b])$, $\alpha, \beta \in \mathbb{R}$, si ha*

$$I(\alpha f + \beta g) = \alpha I(f) + \beta I(g)$$

Anche I_n è una funzione lineare.

Definition 10.3. Sia data $F : C^0([a, b]) \longrightarrow \mathbb{R}$, allora definiamo la norma di F :

$$\|F\| \stackrel{\text{def}}{=} \sup \{ |F(f)| \mid f \in C^0([a, b]), \|f\| = 1 \}$$

dove

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

si verifica che la funzione appena definita è una norma e gode delle seguenti proprietà:

1. $\|F\| = \max \{ |F(f)| \mid f \in C^0([a, b]), \|f\| = 1 \}$
2. $|F(f)| \leq \|F\| \|f\|$

Proposition 10.4.

1. $\|I\| = \int_a^b w(x) dx$
2. $\|I_n\| = \sum_{i=0}^n |w_i|$

Proof: Sia $f \in C^0([a, b])$ t.c. $\|f\| = 1$,

$$\|f\| = \max_{x \in [a, b]} |f(x)| = 1 \Rightarrow |f(x)| \leq 1 \quad \forall x \in [a, b] \quad (1)$$

(1) Dim 1.

$$|I(f)| = \left| \int_a^b w(x) f(x) dx \right| \leq \int_a^b |w(x)| |f(x)| dx = \underbrace{\int_a^b w(x) |f(x)| dx}_{w(x) \geq 0 \quad \forall x} \leq \underbrace{\int_a^b w(x) dx}_{(1)}$$

$$|I(f)| \leq \int_a^b w(x) dx \quad \forall f : \|f\| = 1 \Rightarrow \|I\| = \sup_{\|f\|=1} |I(f)| \leq \int_a^b w(x) dx \quad (2)$$

$$|I(1)| = \left| \int_a^b w(x) dx \right| = \int_a^b w(x) dx$$

$$|I(1)| \leq \sup_{\|f\|=1} |I(f)| = \|I\|$$

$$\Rightarrow \int_a^b w(x) dx \leq \|I\| \quad (3)$$

$$(2), (3) \Rightarrow \|I\| = \int_a^b w(x) dx$$

(2) Dim 2.

Sia $f \in C^0([a, b])$ t.c. $\|f\| = 1$,

$$|I_n(f)| = \left| \sum_{i=0}^n w_i f(x_i) \right| \leq \sum_{i=0}^n |w_i| |f(x_i)| \underbrace{\leq}_{(1)} \sum_{i=0}^n |w_i|$$

$$|I_n(f)| \leq \sum_{i=0}^n |w_i| \quad \forall f : \|f\| = 1 \Rightarrow \|I_n\| \leq \sum_{i=0}^n |w_i| \quad (4)$$

Poiche' $x_0 < x_1 < \dots < x_n$, possiamo definire la seguente funzione:

$$g(x) = \begin{cases} 1 & x = x_i, w_i \geq 0 \\ -1 & x = x_i, w_i < 0 \\ g(x_{i-1}) + \frac{g(x_i) - g(x_{i-1})}{x_i - x_{i-1}}(x - x_{i-1}) & x \in]x_{i-1}, x_i[\end{cases}$$

$|g(x)| \leq 1 \quad \forall x \in [a, b]$, infatti,

$$|g(x_i)| = 1$$

Sia $x \in]x_{i-1}, x_i[\setminus [x_0, \dots, x_n]$

$$x \in]x_{i-1}, x_i[\Rightarrow 0 < x - x_{i-1} < x_i - x_{i-1} \Rightarrow 0 < \frac{x - x_{i-1}}{x_i - x_{i-1}} < 1 \quad (5)$$

$$|g(x)| = \left| g(x_{i-1}) + \frac{g(x_i) - g(x_{i-1})}{x_i - x_{i-1}}(x - x_{i-1}) \right|$$

Se $g(x_i) = g(x_{i-1}) = \pm 1$ allora $|g(x)| = |g(x_{i-1})| = 1$

Se $g(x_i) = -1, g(x_{i-1}) = 1$ allora

$$1 \underbrace{>}_{(5)} g(x) = 1 - 2 \frac{x - x_{i-1}}{x_i - x_{i-1}} \underbrace{>}_{(5)} 1 - 2 = -1 \Rightarrow -1 < g(x) < 1 \Leftrightarrow |g(x)| < 1$$

Analogamente se $g(x_i) = 1, g(x_{i-1}) = -1$ si ha $|g(x)| < 1$

Si verifica facilmente che $g(x)$ e' una funzione continua in $[a, b]$. Nota¹² Quindi,

$$|I_n(g)| = \left| \sum_{i=0}^n w_i g(x_i) \right| \underbrace{=}_{\text{per costruz. di } g} \sum_{i=0}^n |w_i|$$

$$|I_n(g)| \leq \|I_n\|$$

$$\Rightarrow \sum_{i=0}^n |w_i| \leq \|I_n\| \quad (6)$$

$$(4), (6) \Rightarrow \|I_n\| = \sum_{i=0}^n |w_i|$$

□

Definition 10.5. Sia $I_n : C^0([a, b]) \longrightarrow \mathbb{R}$ una formula di quadratura,

$$I_n \text{ e' una formula interpolatoria } \stackrel{\text{def}}{\Leftrightarrow} I_n(f) = I(L_n(f)) \quad \forall f \in C^0([a, b])$$

dove L_n e' il polinomio che interpola f negli $n + 1$ nodi di I_n .

Proposition 10.6. Sia I_n una formula di quadratura interpolatoria. Si ha

$$|I(f) - I_n(f)| \leq \|f - L_n\| \int_a^b w(x) dx$$

¹²In pratica, $g(x)$ e' una spezzata che unisce i punti $(x_i, g(x_i))$

Inoltre, i pesi di I_n sono dati da:

$$w_j = \int_a^b w(x) \lambda_j(x) dx \quad \forall j = 0, \dots, n$$

dove $\lambda_j(x)$ e' dato dalla forma di Lagrange (vedi [6.4,pg.35])

ovvero,

$$I_n(f) = \sum_{j=0}^n w_j f(x_j)$$

Proof: Per semplicita' con $L_n(x)$ indichiamo $L_n(f)(x)$. Abbiamo:

$$\begin{aligned} |I(f) - I_n(f)| &= \left| \int_a^b w(x) f(x) dx - \int_a^b w(x) L_n(x) dx \right| = \left| \int_a^b w(x) (f(x) - L_n(x)) dx \right| \leq \\ &\leq \int_a^b |w(x)| |f(x) - L_n(x)| dx \quad \underbrace{=}_{w(x) \geq 0 \quad \forall x} \int_a^b w(x) |f(x) - L_n(x)| dx \leq \\ &\quad \underbrace{\leq}_{|f(x) - L_n(x)| \leq \|f - L_n\| \quad \forall x} \int_a^b w(x) \|f - L_n\| dx = \|f - L_n\| \int_a^b w(x) dx \end{aligned}$$

Usando la forma di Lagrange (vedi [6.4,pg.35]) abbiamo:

$$\lambda_j(x) = \prod_{\substack{0 \leq i \leq n, \\ i \neq j}} \frac{x - x_i}{x_j - x_i}$$

$$L_n(x) = \sum_{j=0}^n f(x_j) \lambda_j(x)$$

$$I_n(f) = I(L_n) = \int_a^b w(x) L_n(x) dx = \int_a^b w(x) \sum_{j=0}^n f(x_j) \lambda_j(x) dx = \sum_{j=0}^n f(x_j) \int_a^b w(x) \lambda_j(x) dx = \sum_{j=0}^n w_j f(x_j)$$

□

Definition 10.7. Sia data una formula di quadratura $I_n : C^0([a, b]) \longrightarrow \mathbb{R}$,

$$I_n \text{ ha ordine polinomiale } m \stackrel{\text{def}}{\iff} m = \max \{s \in \mathbb{N} \mid I(p) = I_n(p), p \in \mathbb{R}[x]_s\}$$

(Nota¹³) Per indicare che I_n ha ordine polinomiale m useremo la notazione

$$\deg I_n = m$$

Example 10.8. Sia $w(x) \equiv 1$

1. La formula dei trapezi:

$$T(f) \stackrel{\text{def}}{=} \frac{f(a) + f(b)}{2} (b - a)$$

ha ordine polinomiale 1. Infatti, se f e' una retta, allora e' $T(f)$ e' proprio l'area sottesa da tale retta. Il resto e':

$$R = I(f) - T(f) = -\frac{(b-a)^3}{12} f''(\xi), \quad \text{per qualche } \xi \in [a, b]$$

¹³Ricordiamo che $\mathbb{R}[x]_s$ e' l'insieme dei polinomi di $\mathbb{R}[x]$ aventi grado $\leq s$

2. La formula del punto medio:

$$M(f) \stackrel{\text{def}}{=} (b-a)f\left(\frac{b+a}{2}\right)$$

ha ordine polinomiale 1. Basta notare che l'area calcolata da $M(f)$ e' uguale all'area del trapezio calcolata da $T(f)$

3. La formula di Cavalieri-Simpson: si ottiene interpolando f nei punti $(a, f(a)), (b, f(b)), (m, f(m))$, con $m = \frac{a+b}{2}$. Ovvero, se L_n e' il polinomio interpolatore di f ,

$$S(f) \stackrel{\text{def}}{=} I(L_n) = \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]$$

$S(f)$ ha ordine ≥ 2 , infatti, se $p \in \mathbb{R}[x]$ e' di grado due, per l'unicita' dell'interpolazione (vedi [6.3,pg.35]) si ha $L_n = p$ e quindi $S(p) = I(L_n) = I(p)$.

Proposition 10.9.

$$I_n \text{ e' una formula interpolatoria} \Leftrightarrow \deg I_n \geq n$$

Proof: Sia $m = \deg I_n$

(1) Dim $\Rightarrow$

Sia $f(x) \in \mathbb{R}[x]_n$

Per l'unicita' dell'interpolazione (vedi [6.3,pg.35]) abbiamo $L_n(p) = p$

$$I_n(p) \stackrel{\text{def}}{=} I(L_n(p)) = I(p)$$

I_n e' una form. interp.

$$I_n(p) = I(p) \quad \forall p \in \mathbb{R}[x]_n \quad \Rightarrow \quad m \geq n$$

def. di ordine polinom.

(2) Dim $\Leftarrow$

Siano $\lambda_i(x)$ i polinomi di grado n di Lagrange (vedi [6.4,pg.35])

$$I(\lambda_i) \stackrel{\text{def}}{=} I_n(\lambda_i) = \sum_{j=0}^n w_j \lambda_i(x_j) \stackrel{\text{def}}{=} w_i$$

$\deg I_n = m \geq n = \deg \lambda_i$ $\lambda_i(x_j) = \delta_{ji}$

$$I_n(f) = \sum_{j=0}^n w_j f(x_j) = \sum_{j=0}^n I(\lambda_j) f(x_j) \stackrel{\text{linearita'}}{=} I\left(\sum_{j=0}^n \lambda_j(x) f(x_j)\right) \stackrel{[6.4,pg.35]}{=} I(L_n(f))$$

$$I_n(f) = I(L_n(f)) \quad \forall f \in C^0([a, b]) \Rightarrow I_n \text{ e' interpolatoria}$$

□

Proposition 10.10. Sia I_n una formula interpolatoria. Conoscendo i nodi x_i , $i = 0, \dots, n$ si possono calcolare i pesi risolvendo il seguente sistema $(n+1) \times (n+1)$:

$$\int_a^b w(x) x^i dx = \sum_{j=0}^n w_j x_j^i \quad i = 0, \dots, n$$

nelle incognite w_j $j = 0, \dots, n$.

Proof: Per la prop [10.9,pg.72] si ha $\deg I_n \geq n$, quindi I_n e' esatta per i polinomi di grado $\leq n$, quindi

$$I(x^i) = I_n(x^i) \quad \forall i = 0, \dots, n \Leftrightarrow \int_a^b w(x)x^i dx = \sum_{j=0}^n w_j x_j^i \quad i = 0, \dots, n$$

□

Theorem 10.11. *Sia*

1. $f \in C^0([a, b])$
2. $\{I_n\}$ una successione di formule di quadratura qualsiasi
3. $\lim_{n \rightarrow +\infty} \deg I_n = \infty$
4. $\exists M > 0 : \|I_n\| \leq M \quad \forall n \geq 0$

allora

$$\lim_{n \rightarrow +\infty} I_n(f) = I(f)$$

Proof: Sia $m_n = \deg I_n$. Sia $\varphi_{m_n} \in \mathbb{R}[x]_{m_n}$ il polinomio dato dal thm [7.4,pg.44], relativamente alla funzione f .

$$\begin{aligned} \deg \varphi_{m_n} \leq m_n = \deg I_n &\Rightarrow I(\varphi_{m_n}) = I_n(\varphi_{m_n}) \\ |I_n(f) - I(f)| &= |I_n(f) - I_n(\varphi_{m_n}) + I(\varphi_{m_n}) - I(f)| \leq |I_n(f) - I_n(\varphi_{m_n})| + |I(\varphi_{m_n}) - I(f)| = \\ &\underbrace{=}_{\text{linearita'}} |I_n(f - \varphi_{m_n})| + |I(\varphi_{m_n} - f)| \leq \underbrace{\|I_n\| \|f - \varphi_{m_n}\| + \|I\| \|\varphi_{m_n} - f\|}_{[10.3,pg.69]} = \|f - \varphi_{m_n}\| (\|I\| + \|I_n\|) \quad (1) \end{aligned}$$

$$\lim_{n \rightarrow +\infty} m_n = \lim_{n \rightarrow +\infty} \deg I_n = \infty \xRightarrow{[7.4,pg.44]} \lim_{n \rightarrow +\infty} \varphi_{m_n} = f \Rightarrow \lim_{n \rightarrow +\infty} \|f - \varphi_{m_n}\| = 0 \quad (2)$$

Per Hp la successione $\|I_n\|$ e' limitata, $\|I\| \in \mathbb{R} \Rightarrow \|I\| + \|I_n\|$ e' limitata (3)

$$(1), (2), (3) \Rightarrow \lim_{n \rightarrow +\infty} |I_n(f) - I(f)| \leq \lim_{n \rightarrow +\infty} \|f - \varphi_{m_n}\| (\|I\| + \|I_n\|) = 0 \Rightarrow \lim_{n \rightarrow +\infty} I_n(f) = I(f)$$

□

Proposition 10.12. *Calcolo pratico di I_n .*

Consideriamo una formula di quadratura $I_n : C^0([a, b]) \rightarrow \mathbb{R}$ che approssimi bene la funzione I .

Per il calcolo di $I_n(f_i)$, $f_i \in C^0([a, b]) \quad \forall i = 1, \dots, m$, si puo' adottare il seguente algoritmo:

1. *Si sceglie un intervallo $[\alpha, \beta] \subseteq [a, b]$ conveniente (ad esempio $[0, 1]$).*
2. *Si scelgono $n + 1$ nodi in $[\alpha, \beta]$.*
3. *Si pre-computano i pesi w_i relativamente all'intervallo $[\alpha, \beta]$. Questo e' il calcolo piu' pesante, in quanto richiede il calcolo di $\int_{\alpha}^{\beta} w(x)$.*
Nota: per una formula interpolatoria, i pesi w_i dipendono solo dai nodi (vedi [10.6,pg.70])
4. *Per il calcolo di $I_n(f_i)$:*

(a) *Si calcolano i nuovi pesi e nodi:*

$$\begin{aligned} w'_j &= w_j \frac{b-a}{\beta-\alpha} \\ x'_j &= \frac{b-a}{\beta-\alpha}(x_j - \alpha) + a \end{aligned}$$

(b) Si ha la seguente approssimazione:

$$I_n(f_i) \approx \sum_{j=0}^n w'_j f_i(x'_j)$$

Proof: Per ipotesi:

$$\int_{\alpha}^{\beta} f(\xi) d\xi \approx I_n(f) = \sum_{j=0}^n w_j f(x_j) \quad (1)$$

Sia $f = f_i$. Effettuiamo il seguente cambio di variabili:

$$x = \varphi(\xi) = \frac{b-a}{\beta-\alpha}(\xi - \alpha) + a, \quad \xi \in [\alpha, \beta], \quad x \in [a, b]$$

in pratica stiamo mappando l'intervallo $[\alpha, \beta]$ in $[a, b]$, imponendo che $\xi - \alpha : \beta - \alpha = x - a : b - a$

$$\varphi^{-1}(x) = \frac{\beta - \alpha}{b - a}(x - a) + \alpha$$

$$\begin{aligned} \int_a^b f(x) dx &= \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} \varphi'(\xi) f(\varphi(\xi)) d\xi = \\ &= \int_{\alpha}^{\beta} \frac{b-a}{\beta-\alpha} f\left(\frac{b-a}{\beta-\alpha}(\xi - \alpha) + a\right) d\xi \underset{(1)}{\approx} I_n\left(\frac{b-a}{\beta-\alpha} f\left(\frac{b-a}{\beta-\alpha}(\xi - \alpha) + a\right)\right) = \\ &= \sum_{j=0}^n w_j \left[\frac{b-a}{\beta-\alpha} f\left(\frac{b-a}{\beta-\alpha}(x_j - \alpha) + a\right) \right] = \sum_{j=0}^n w'_j f(x'_j) \end{aligned}$$

□

Proposition 10.13. *Formule di Newton-Cotes.*

Le formule di Newton-Cotes sono formule interpolatorie ottenute scegliendo una distribuzione uniforme dei nodi. Si classificano in formule chiuse e formule aperte

Formule chiuse Gli estremi a, b sono dei nodi: $x_0 = a, x_n = b$,

$$\begin{aligned} h &= \frac{b-a}{n} \\ x_i &= a + ih \quad i = 0, \dots, n \end{aligned}$$

Formule aperte Gli estremi non sono dei nodi:

$$\begin{aligned} h &= \frac{b-a}{n+2} \\ x_i &= a + (i+1)h \quad i = 0, \dots, n \end{aligned}$$

Example 10.14. Chiusa Per $n = 1$ si ottiene la formula del trapezio: infatti, stiamo interpolando la funzione in 2 punti, ovvero negli estremi. Il polinomio interpolatore e' la retta che passa per i punti $(a, f(a)), (b, f(b))$. $I_n(f)$ e' l'area sottostante alla retta.

Analogamente, per $n = 2$ si ottiene la formula di Simpson.

Aperta Per $n = 0$ si ottiene la formula del punto medio.

Proposition 10.15. Consideriamo una qualsiasi successione $\{I_n\}$ di formule di Newton-Cotes, allora si dimostra che

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^n |w_i| = \infty \quad \underbrace{\Leftrightarrow}_{[10.4, \text{pg. } 69]} \lim_{n \rightarrow +\infty} \|I_n\| = \infty$$

Per questo motivo, non possiamo applicare il thm [10.11, pg. 73]. Quindi, non abbiamo alcuna garanzia sulla convergenza delle formule di Newton-Cotes.

Proposition 10.16. Consideriamo una formula I_n di Newton-Cotes. Si dimostra che

$$\deg I_n = \begin{cases} n & n \text{ dispari} \\ n+1 & n \text{ pari} \end{cases}$$

Posto $R_n(f) = I(f) - I_n(f)$, chiamato resto della formula di quadratura, si dimostra che:

n pari:

$$R_n(f) = K \frac{f^{(n+2)}(\xi)}{(n+2)!} h^{n+3}, \quad \text{per qualche } \xi \in [a, b]$$

$$K = \begin{cases} \int_{-1}^{n+1} x r_n(x) dx & I_n \text{ formula aperta} \\ \int_0^n x r_n(x) dx & I_n \text{ formula chiusa} \end{cases}$$

n dispari:

$$R_n(f) = K \frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+2}, \quad \text{per qualche } \xi \in [a, b]$$

$$K = \begin{cases} \int_{-1}^{n+1} r_n(x) dx & I_n \text{ formula aperta} \\ \int_0^n r_n(x) dx & I_n \text{ formula chiusa} \end{cases}$$

dove

$$r_n(x) = \frac{x!}{(x-n-1)!} = x(x-1)(x-2)\dots(x-n)$$

In particolare, per $n = 0, \dots, 4$, R_n assume i seguenti valori:

n	Chiusa	Aperta
0		$\frac{1}{3}h^3 f^{(2)}(\xi)$
1	$-\frac{1}{12}h^3 f^{(2)}(\xi)$	$\frac{3}{4}h^3 f^{(2)}(\xi)$
2	$-\frac{1}{90}h^5 f^{(4)}(\xi)$	$\frac{14}{45}h^5 f^{(4)}(\xi)$
3	$-\frac{3}{80}h^5 f^{(4)}(\xi)$	$\frac{95}{144}h^5 f^{(4)}(\xi)$
4	$-\frac{8}{945}h^7 f^{(6)}(\xi)$	$\frac{41}{140}h^7 f^{(6)}(\xi)$

Proposition 10.17.

$$\max \{ \deg I_n \mid I_n \text{ formula di quadratura su } n+1 \text{ pesi e nodi} \} = 2n+1$$

ovvero, l'ordine polinomiale massimo raggiungibile da una formula di quadratura I_n e' $2n+1$.
In particolare, se I_n e' una formula di quadratura interpolatoria sui nodi $x_0, \dots, x_n \in [a, b]$ e pesi $w_0, \dots, w_n$, si ha

$$q_n \perp f \quad \forall f \in \mathbb{R}[x]_n \Rightarrow \deg I_n = 2n+1$$

dove

$$q_n(x) = \prod_{i=0}^n (x - x_i)$$

$$\text{Il prodotto scalare usato e': } f \cdot g = I(fg) \quad \forall f, g \in C^0([a, b])$$

Proof:

(1) Dim. che $\deg I_n < 2n+2$ per qualsiasi scelta dei pesi e dei nodi

Dimostriamo che esiste un polinomio p di grado $2n+2$ per cui $I_n(f) \neq I(f)$.

I_n e' una formula di quadratura, quindi

$$I_n(f) = \sum_{i=0}^n w_i f(x_i)$$

Poniamo

$$q(x) = \prod_{i=0}^n (x - x_i)$$

$$p(x) = q(x)^2$$

$$\deg q(x) = n+1, \quad \deg p(x) = 2n+2$$

si ha

$$q(x_i) = p(x_i) = 0 \quad \forall i = 0, \dots, n$$

$$I(p) = \int_a^b w(x)p(x) dx \quad \underbrace{\qquad}_{p(x), w(x) \geq 0 \quad \forall x \in [a, b], \quad \exists \xi \in [a, b]: w(\xi)p(\xi) > 0, \quad [\text{AII}, 1.22, \text{pg.17}]} > 0$$

$$I_n(p) = \sum_{i=0}^n w_i p(x_i) = 0$$

$$\Rightarrow I_n(p) \neq I(p)$$

(2) Dim 2.

Sia $f \in \mathbb{R}[x]_{2n+1}$. Dimostriamo che $I(f) = I_n(f)$.

Sia $L_n(f)$ il polinomio che interpola f nei punti $x_0, \dots, x_n$. Poiche' I_n e' interpolatoria, abbiamo

$$I_n(f) = I(L_n(f))$$

Per il thm [6.8,pg.37] abbiamo

$$f(x) = L_n(f)(x) + \underbrace{f[x_0, \dots, x_n, x]}_{g(x)} q_n(x)$$

$$f \in \mathbb{R}[x]_{2n+1}, \quad q_n \in \mathbb{R}[x]_{n+1}, \quad L_n(f) \in \mathbb{R}[x]_n \Rightarrow g(x) = f[x_0, \dots, x_n, x] \in \mathbb{R}[x]_n$$

$$q_n \perp h \quad \forall h \in \mathbb{R}[x]_n \Rightarrow q_n \perp g \Leftrightarrow I(gq_n) = 0$$

$$I(f) = I(L_n(f)) + I(gq_n) \quad \underbrace{\qquad}_{I(gq_n)=0} = I(L_n(f)) = I_n(f)$$

□

Definition 10.18. Sia I_n una formula di quadratura,

$$I_n \text{ e' gaussiana} \stackrel{\text{def}}{\Leftrightarrow} \deg I_n = 2n + 1$$

Proposition 10.19. Metodo per il calcolo dei nodi $x_0, \dots, x_n$ che rendono la formula interpolatoria I_n una gaussiana.

Sia $q_n(x) = \prod_{i=0}^n x - x_i$

1. $q_n(x) = a_0 + a_1x + \dots + x^{n+1}$ (Nota: ¹⁴)
2. Si risolve il sistema lineare $(n+1) \times (n+1)$

$$\int_a^b w(x) x^k q_n(x) dx = 0 \quad \forall k = 0, \dots, n$$

nelle incognite $a_0, \dots, a_n$. (Nota¹⁵)

3. Si trovano gli zeri di $q_n(x)$. Tali zeri sono i nodi $x_0, \dots, x_n$ cercati.

Nota: Per calcolare i nodi in piu' intervalli $[a_1, b_1], \dots, [a_r, b_r]$ convien pre-computare $x_0, \dots, x_n$ in un intervallo fisso (ad esempio $[0, 1]$) e usare la prop [10.12, pg. 73].

Proof:

$$\int_a^b w(x) x^k q_n(x) dx = 0 \Leftrightarrow I(x^k q_n) = 0 \Leftrightarrow x^k \cdot q_n = 0 \Leftrightarrow x^k \perp q_n \quad \forall k = 0, 1, \dots, n$$

Poiche' $1, x, \dots, x^n$ e' una base di $\mathbb{R}[x]_n$, si ha

$$q_n \perp x^k \quad \forall k = 0, \dots, n \quad \Leftrightarrow \quad q_n \perp f \quad \forall f \in \mathbb{R}[x]_n$$

[GII, 2.23, pg. 40]

Quindi per la prop. [10.17, pg. 76] abbiamo la tesi. □

Example 10.20. 1. Per $w(x) \equiv 1$, $[a, b] = [-1, 1]$, $\deg I_n = 2n + 1$, I_n viene chiamata *formula di Gauss-Legendre*

2. Per $w(x) \equiv \frac{1}{\sqrt{1-x^2}}$, $[a, b] = [-1, 1]$, $\deg I_n = 2n + 1$, I_n viene chiamata *formula di Gauss-Chebyshev*. Si ha:

$$w_i = \frac{\pi}{n+1} \quad \forall i = 0, \dots, n$$

I nodi x_i sono le radici di $t_{n+1}(x)$. Vedi [6.10, pg. 40]

10.1 Formule composite

Definition 10.21. Suddividiamo l'intervallo $[a, b]$ in N intervallini $[x_{i-1}, x_i]$ $i = 1, \dots, N$.

Possiamo allora spezzare il calcolo dell'integrale:

$$\int_a^b f(x) dx = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} f(x) dx = \sum_{i=1}^N I^{(i)}(f) dx \approx \sum_{i=1}^N I_n^{(i)}(f) \stackrel{\text{def}}{=} C_N(f)$$

$C_N(f)$ si chiama formula di *quadratura composita*.

¹⁴ $q_n(x)$ e' monico

¹⁵Stiamo supponendo $w(x)$ si possa integrare facilmente (ad esempio, $w(x)$ puo' essere un polinomio). Poiche' $x^k q_n(x)$ e' un polinomio, il sistema diventa lineare.

Proposition 10.22. *Formule composite:*

1. *Formula dei trapezi:*

$$C_N(f) = \sum_{i=1}^N \frac{f(x_{i-1}) + f(x_i)}{2} (x_i - x_{i-1})$$

Se $x_i - x_{i-1} = h = \frac{b-a}{N}$, allora

$$C_N(f) = h \left(\frac{f(a) + f(b)}{2} + \sum_{i=1}^{N-1} f(a + ih) \right)$$

Proof:

1. Trapezi uniforme:

$$\begin{aligned} C_N(f) &= \sum_{i=1}^N \frac{f(x_{i-1}) + f(x_i)}{2} h = \frac{h}{2} \sum_{i=1}^N (f(x_{i-1}) + f(x_i)) \\ &= \sum_{i=1}^N (f(x_{i-1}) + f(x_i)) = \sum_{i=1}^N f(x_{i-1}) + \sum_{i=1}^N f(x_i) = \\ &= \sum_{i=0}^{N-1} f(x_i) + \sum_{i=1}^N f(x_i) = f(x_0) + f(x_N) + \sum_{i=1}^{N-1} f(x_i) + \sum_{i=1}^{N-1} f(x_i) = f(x_0) + f(x_N) + 2 \sum_{i=1}^{N-1} f(x_i) \\ C_N(f) &= \frac{h}{2} \left(f(x_0) + f(x_N) + 2 \sum_{i=1}^{N-1} f(x_i) \right) = h \left(\frac{f(x_0) + f(x_N)}{2} + \sum_{i=1}^{N-1} f(x_i) \right) \\ \text{Poiche' } x_0 &= a, x_N = b, x_i = a + ih \quad \forall i = 0, \dots, N, \text{ abbiamo} \\ C_N(f) &= h \left(\frac{f(a) + f(b)}{2} + \sum_{i=1}^{N-1} f(a + ih) \right) \end{aligned}$$

□

Example 10.23. Per la formula dei trapezi composita:

$$R = -\frac{b-a}{12} h^2 f''(\xi)$$

dove $[a, b]$ e' stato suddiviso uniformemente in N intervallini, con $h = \frac{b-a}{N}$.

Proposition 10.24. *Estrapolazione di Richardson.*

Sia $\{I_n(f, h)\}_{h \in \mathbb{R}}$ una famiglia di formule di quadratura. Supponiamo che il resto di $I_n(f, h)$ sia del tipo:

$$\begin{aligned} R_n(f, h) &= c_1 h^{p_1} + c_2 h^{p_2} + \dots + c_k h^{p_k} + \dots \\ p_1 &< p_2 < \dots < p_k < \dots \end{aligned}$$

Scelto $q > 1$, possiamo costruire la seguente successione di formule interpolatorie:

$$\begin{aligned} F_1(f, h) &= I_n(f, h) \\ F_{k+1}(f, h) &= \frac{q^{p_k} F_k(f, h) - F_k(f, qh)}{q^{p_k} - 1} \end{aligned}$$

Quindi se $h < 1$, il resto di F_{k+1} tende a 0 (per $k \rightarrow \infty$) esponenzialmente. In altre parole, usando poche iterazioni dell'extrapolazione di Richardson, miglioriamo notevolmente l'accuratezza della formula di quadratura.

$$\begin{array}{cccc} F_1(f, h) & & & \\ & F_2(f, h) & & \\ F_1(f, qh) & & F_3(f, h) & \\ & F_2(f, qh) & & F_4(f, h) \\ F_1(f, q^2h) & & F_3(f, qh) & \\ & F_2(f, q^2h) & & \\ F_1(f, q^3h) & & & \end{array}$$
$$\begin{aligned}
I(f) &= I_n(f, h) + R_n(f, h) = I_n(f) + c_1 h^{p_1} + O(h^{p_2}) \\
I(f) &= F_1(f, h) + c_1 h^{p_1} + O(h^{p_2}) \quad (1) \\
I(f) &= F_1(f, qh) + c_1 (qh)^{p_1} + O(h^{p_2}) \quad (2) \\
(1) \quad &\underbrace{\Rightarrow}_{\text{multiplicando per } q^{p_1}} \quad q^{p_1} I(f) = q^{p_1} F_1(f, h) + c_1 (qh)^{p_1} + O(h^{p_2}) \quad (1.1)
\end{aligned}$$

Procedendo per induzione si ha la tesi.

$$x_i - x_{i-1} = h = \frac{b-a}{N} \quad \forall i = 1, \dots, N$$

$$C_N(f) = h \left(\frac{f(a) + f(b)}{2} + \sum_{i=1}^{N-1} f(a + ih) \right)$$

quindi e' possibile applicare l'estrapolazione di Richardson ad f (vedi [10.24,pg.78])

Sia $f \in C^0([a, b])$, $I_n(a, b)$ una formula di quadratura che approssima $I(f)$ sull'intervallo a, b . E sia $I'_n(a, b)$ una approssimazione migliore di $I_n(a, b)$ (ad esempio ottenuta tramite l'estrapolazione,

vedi [10.24,pg.78]).

L'integrazione adattiva si puo' descrivere semplicemente dall'algoritmo stesso:

IntAdattiva(α, β) :

$M =: I_n(\alpha, \beta)$

$N =: I'_n(\alpha, \beta)$

Se $|M - N| < \frac{\varepsilon}{b-a}(\beta - \alpha)$

$I =: N$

Altrimenti :

Suddividiamo l'intervallo e calcolando ricorsivamente l'integrale composito:

$I =: \text{IntAdattiva}\left(\alpha, \frac{\alpha + \beta}{2}\right) + \text{IntAdattiva}\left(\frac{\alpha + \beta}{2}, \beta\right)$

$I =: \text{IntAdattiva}(a, b)$

il risultato I e' un approssimazione di $I(f)$ con un errore $< \varepsilon$.

(!)Bisogna pero' fare attenzione sulla scelta di I'_n (vedi dimostrazione)(!)

Proof: L'integrazione adattiva al k -esimo passo utilizza la seguente formula composita:

$$C_{2^k} = \sum_{i=1}^{2^k} I_n(x_{i-1}, x_i)$$

$$\text{con } h = \frac{b-a}{2^k}, \quad x_i = a + ih$$

Sia $I^{(i)} = I^{(i)}(f)$ l'integrale esatto di f in $[x_{i-1}, x_i]$.

(!)Se I'_n e' sufficientemente accurato(!) abbiamo:

$$|I^{(i)} - I_n(x_{i-1}, x_i)| \approx |I'_n(x_{i-1}, x_i) - I_n(x_{i-1}, x_i)| = |M - N| < \frac{\varepsilon}{b-a}(x_i - x_{i-1})$$

Quindi

$$\begin{aligned} |I(f) - C_{2^k}| &= \left| \sum_{i=1}^{2^k} I^{(i)} - I_n(x_{i-1}, x_i) \right| \leq \sum_{i=1}^{2^k} |I^{(i)} - I_n(x_{i-1}, x_i)| < \\ &< \sum_{i=1}^{2^k} \frac{\varepsilon}{b-a}(x_i - x_{i-1}) \underbrace{\quad}_{\text{somma telescopica}} \frac{\varepsilon}{b-a}(b-a) = \varepsilon \end{aligned}$$

□

11 Formulario

Sistemi lineari

1. Norme:

$$\rho(A) = \lambda_M(A) = \max\{|\lambda| \mid \lambda \text{ autoval } A\}$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{ij}|$$

$$\|A\|_1 = \|{}^t A\|_\infty$$

$$\|A\|_2 = \sqrt{\lambda_M({}^t A A)}$$

$$\|A^{-1}\|_2 = \frac{1}{\sqrt{\lambda_m({}^t A A)}}$$

$$A = {}^t A \Rightarrow \|A\|_2 = |\lambda_M(A)|$$

$$A = {}^t A \Rightarrow \|A^{-1}\|_2 = \frac{1}{|\lambda_m(A)|}$$

$$\rho(A) = \inf_{p,q \in \mathbb{N}^*} \|A\|_{p,q}$$

2. Triangolare superiore:

$$x_i = \frac{b_i - \sum_{j=i+1}^n a_{ij} x_j}{a_{ii}} \quad \forall i = n, \dots, 1$$

Triangolare inferiore:

$$x_i = \frac{b_i - \sum_{j=1}^{i-1} a_{ij} x_j}{a_{ii}} \quad \forall i = 1, \dots, n$$

3. Condizionamento:

$$A(x + \Delta x) = b + \Delta b$$

$$\frac{\|\Delta x\|}{\|x\|} \leq \mu(A) \frac{\|\Delta b\|}{\|b\|}$$

$$\mu_2(A) = \frac{\sqrt{\lambda_M({}^t A A)}}{\sqrt{\lambda_m({}^t A A)}}$$

$$A = {}^t A \Rightarrow \mu_2(A) = \frac{|\lambda_M|}{|\lambda_m|}$$

4. Gauss (pivot):

pivot_riga k :

$$\text{Sia } \bar{i} \text{ t.c. } |a_{\bar{i}k}| := \max_{i=k, \dots, n} |a_{ik}|$$

scambio_righe $\bar{i}, k$

pivot_scalato k :

$$d_i := \frac{|a_{ik}|}{\max_{j=k, \dots, n} |a_{ij}|}$$

$$\text{Sia } \bar{i} \text{ t.c. } d_{\bar{i}} = \max_{i=1, \dots, n} d_i$$

$$\forall k = 1, \dots, n-1$$

pivot k

$$\forall i = k+1, \dots, n$$

$$a_{ik} := \frac{a_{ik}}{a_{kk}}$$

$$\forall j = k+1, \dots, n+1$$

$$a_{ij} := a_{ij} - a_{ik} a_{kj}$$

5. Fattorizzazione Cholesky

$A \in \mathbb{R}^{n \times n}$, simmetrica, definita pos.

$A = R^t R$, $R \in \mathbb{R}^{n \times n}$ triang. inf.

$$a_{ij} = \sum_{k=1}^{\min\{i,j\}} R_{ik} R_{jk} \quad \forall i, j$$

$$\forall j = 1, \dots, n$$

$$r_{jj} := \sqrt{a_{jj} - \sum_{k=1}^{j-1} r_{jk}^2}$$

$$\forall i = j+1, \dots, n$$

$$r_{ij} := \frac{a_{ij} - \sum_{k=1}^{j-1} r_{ik} r_{jk}}{r_{jj}}$$

6. Gershgorin-Hadamard

$$A \in \mathbb{C}^{n \times n}$$

$$D_i = \{z \in \mathbb{C} \mid |z - a_{ii}| \leq R_i\}$$

$$R_i = \sum_{j \neq i} |a_{ij}|$$

$$\{\text{autoval. di } A\} \subseteq \bigcup_{i=1}^n D_i$$

7. Metodo delle potenze:

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$$

$$x_0 \in \mathbb{C}^{n,1}, \quad x_0 \neq 0$$

$$y_k = \frac{x_k}{\|x_k\|}, \quad x_{k+1} = A y_k$$

$$\sigma_k = \frac{{}^t x_k A x_k}{{}^t x_k x_k} = \frac{{}^t y_k x_{k+1}}{{}^t y_k y_k}$$

$$z_0 = x_0, \quad z_{k+1} = A z_k, \quad \sigma'_k = \frac{{}^t z_k A z_k}{{}^t z_k z_k}$$

$$\lim_{k \rightarrow \infty} \sigma_k = \lim_{k \rightarrow \infty} \sigma'_k = \lambda_1$$

$$\text{Crit. arresto: } \frac{\|(A - \sigma_k I) y_k\|}{\sigma_k \|y_k\|} < \varepsilon$$

$$\text{oppure } |\sigma_{k+1} - \sigma_k| < \varepsilon$$

Interpolazione

1. Lagrange:

$$\lambda_j(x) = \prod_{0 \leq i \leq n, i \neq j} \frac{x - x_i}{x_j - x_i}$$

$$p(x) = \sum_{j=0}^n y_j \lambda_j(x)$$

$$\sum_{j=0}^n \lambda_j(x) (x_j - x)^k \equiv \begin{cases} 1 & k=0 \\ 0 & k>0 \end{cases}$$

$$\lambda_j(x_i) = \delta_{ij}$$

2. Differenze divise:

$$f[x_0] = f(x_0)$$

$$f[x_0, \dots, x_j] =$$

$$= \frac{f[x_1, \dots, x_j] - f[x_0, \dots, x_{j-1}]}{x_j - x_0}$$

$$f[x_0] = f(x_0)$$

$$f[x_1] = f(x_1) \quad f[x_0, x_1]$$

$$f[x_2] = f(x_2) \quad f[x_1, x_2] \quad f[x_0, x_1, x_2]$$

3. Newton:

$$w_0(x) = 1, \quad w_j(x) = \prod_{i=0}^{j-1} (x - x_i)$$

$$p_n(x) = \sum_{i=0}^n f[x_0, \dots, x_i] w_i(x)$$

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x)$$

4. Chebyshev

$$t_n(x) = \cos n\varphi(x) : [-1, 1] \longrightarrow \mathbb{R}$$

$$\varphi(x) = \arccos x : [-1, 1] \longrightarrow [0, \pi]$$

$$t_0(x) = 1, \quad t_1(x) = x$$

$$t_n(x) = 2t_{n-1}(x)x - t_{n-2}(x) \quad \forall n > 1$$

$$\deg t_n(x) = n \quad \forall n \geq 0$$

$$c_n = 2^{n-1} \quad \forall n \geq 1$$

radici di t_n :

$$\alpha_k = \cos\left(\frac{\pi}{2n} + k\frac{\pi}{n}\right) \quad k = 0, \dots, n-1$$

$$\text{max: } M_k = \cos\left(k\frac{2\pi}{n}\right) \quad k = 0, \dots, n$$

$$\text{min: } m_k = \cos\left(k\frac{2\pi}{n} + \frac{\pi}{n}\right) \quad k = 0, \dots, n$$

5. Polinomio di Bernstein:

$$B_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}$$

$$\sum_{k=0}^n B_{n,k}(x) = 1$$

$$B_n(f)(x) = \sum_{k=0}^n f\left(\frac{k}{n}\right) B_{n,k}(x)$$

$$\lim_{n \rightarrow +\infty} B_n(f) \stackrel{u.}{=} f \text{ in } [0, 1]$$

$$\text{dove } f \in C^0([0, 1])$$

6. Minimi quadrati

$$\varphi(x, \mathbf{a}) = \sum_{j=1}^m a_j \varphi_j(x), \quad \phi \in \mathbb{R}^{n \times m}$$

$$\phi_{jk} = \varphi_k(x_j), \quad {}^t \phi \phi \mathbf{a} = {}^t \phi \mathbf{y}, \quad \det {}^t \phi \phi \neq 0$$

$$\varepsilon(\mathbf{v})_i = y_i - \varphi(x_i, \mathbf{v}), \quad \|\varepsilon(\mathbf{a})\|_2^2 \text{ minima}$$

Zeri e punti fissi

1.

$$\text{Newton: } x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Secanti:

$$x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}$$

Bisezione:

While :

$$c := \frac{a+b}{2}$$

$$\text{Se } |b-a| < 2\varepsilon \vee f(c) = 0 :$$

break

$$\text{Se } f(c)f(a) > 0 : \quad a := c$$

$$\text{Se } f(c)f(a) < 0 : \quad b := c$$

$$\alpha := c$$

2. Criterio di arresto di Aitken per successioni lineari:

$$\text{Hp: } x_n = \varphi(x_{n-1}) \rightarrow \alpha = \varphi(\alpha)$$

$$\text{Hp: } \varphi'(\alpha) \neq 0, \quad x_n \neq \alpha \quad \forall n$$

$$\left| x_n - \frac{x_{n+1}x_{n-1} - x_n^2}{x_{n+1} + x_{n-1} - 2x_n} \right| < \varepsilon$$

3. Minimo di f unimodale:

While

$$x_1 := a + (1-r)(b-a)$$

$$x_2 := a + r(b-a)$$

$$\text{Se } f(x_1) > f(x_2) : \quad a := x_1$$

$$\text{Altrimenti:} \quad b := x_2$$

$$\text{Se } b-a < \varepsilon :$$

$$x_* := a$$

break

Integrali

1. Norma

$$\|I\| = \int_a^b w(x) dx, \quad \|I_n\| = \sum_{i=0}^n |w_i|$$

2. Formula interpolatoria I_n :

$$\deg I_n \geq n$$

$$|I(f) - I_n(f)| \leq \|f - L_n\| \int_a^b w(x) dx$$

$$w_j = \int_a^b w(x) \lambda_j(x) dx \quad \forall j = 0, \dots, n$$

Sistema per il calcolo dei pesi:

$$\int_a^b w(x) x^i dx = \sum_{j=0}^n w_j x_j^i \quad \forall i = 0, \dots, n$$

3. Cambio di pesi e nodi:

$$x_j, w_j \text{ su } [\alpha, \beta], \quad x'_j, w'_j \text{ su } [a, b]$$

$$w'_j = w_j \frac{b-a}{\beta-\alpha}$$

$$x'_j = \frac{b-a}{\beta-\alpha} (x_j - \alpha) + a$$

4. Newton-Cotes: formula interpolatoria su $n+1$ nodi uniformi.

Chiusa

$$x_0 = a, \quad x_n = b, \quad h = \frac{b-a}{n},$$

$$x_i = a + ih \quad i = 0, \dots, n$$

Aperta

$$h = \frac{b-a}{n+2}, \quad x_i = a + (i+1)h \quad \forall i = 0, \dots, n$$

$$\deg I_n = \begin{cases} n & n \text{ dispari} \\ n+1 & n \text{ pari} \end{cases}$$

Resto: n	Chiusa	Aperta
0		$\frac{1}{3} h^3 f^{(2)}(\xi)$
1	$-\frac{1}{12} h^3 f^{(2)}(\xi)$	$\frac{3}{4} h^3 f^{(2)}(\xi)$
2	$-\frac{1}{90} h^5 f^{(4)}(\xi)$	$\frac{14}{45} h^5 f^{(4)}(\xi)$
3	$-\frac{3}{80} h^5 f^{(4)}(\xi)$	$\frac{95}{144} h^5 f^{(4)}(\xi)$
4	$-\frac{8}{945} h^7 f^{(6)}(\xi)$	$\frac{41}{140} h^7 f^{(6)}(\xi)$

5. Formula dei trapezi

$$T(f) \stackrel{\text{def}}{=} \frac{f(a) + f(b)}{2} (b-a)$$

$$w(x) \equiv 1, \quad \deg T = 1$$

$$\text{Resto: } -\frac{(b-a)^3}{12} f''(\xi), \quad \xi \in [a, b]$$

Composita, con distribuzione uniforme dei nodi:

$$C_N(f) = h \left(\frac{f(a) + f(b)}{2} + \sum_{i=1}^{N-1} f(a+ih) \right)$$

$$x_i - x_{i-1} = h = \frac{b-a}{N}$$

$$\text{Resto: } -\frac{b-a}{12} h^2 f''(\xi)$$

Resto di Eulero-MacLaurin:

$$c_1 h^2 + c_2 h^4 + \dots + c_k h^{2k} + \dots$$

6. Cavalieri-Simpson (formula interpolatoria)

$$\text{Nodi } x_0 = a, x_1 = \frac{a+b}{2}, x_2 = b$$

$$S(f) = \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]$$

$$w(x) \equiv 1, \quad \deg S = 2$$

$$\text{Resto formula composita: } -\frac{b-a}{2880} h^4 f^{(4)}(\xi)$$

$$\text{con } x_i - x_{i-1} = h = \frac{b-a}{N}$$

7. Gauss-Chebyshev:

$$w(x) \equiv \frac{1}{\sqrt{1-x^2}}, \quad [a, b] = [-1, 1], \quad \deg I_n = 2n+1$$

$$w_i = \frac{\pi}{n+1} \quad \forall i = 0, \dots, n$$

$$x_i = \cos \left[\frac{1}{n+1} \left(\frac{\pi}{2} + i\pi \right) \right] \quad i = 0, \dots, n$$

8. Estrapolazione di Richardson:

$$F_1(f, h) = I_n(f, h), \quad h < 1$$

$$F_{k+1}(f, h) = \frac{q^{pk} F_k(f, h) - F_k(f, qh)}{q^{pk} - 1}$$

Per la formula dei trapezi: $p_k = 2k$

9. Integrazione adattiva:

IntAdattiva(α, β) :

$$M =: I_n(\alpha, \beta)$$

$$N =: I'_n(\alpha, \beta)$$

$$\text{Se } |M - N| < \frac{\varepsilon}{b-a} (\beta - \alpha)$$

$$I =: N$$

Altrimenti :

$$I =: \text{IntAdattiva} \left(\alpha, \frac{\alpha + \beta}{2} \right) + \text{IntAdattiva} \left(\frac{\alpha + \beta}{2}, \beta \right)$$

$$I =: \text{IntAdattiva}(a, b)$$

12 *

Index

*, 81

Applicazioni, 24

Approssimazione polinomiale, 43

Autovalori, 27

Criteri di arresto, 63

differenza divisa, 36

Equazioni non lineari, 53

estrapolazione di Richardson, 78

Fattorizzazione di Cholesky, 26

Fattorizzazione LU, 19

flops, 2

formula di Gauss-Chebyshev, 77

formula di Gauss-Legendre, 77

formula di quadratura, 68

formula quadratura interpolatoria, 70

formula Simpson, 71

formula trapezi, 71

Formule composite, 77

Integrali, 68

Interpolazione e approssimazione, 33

Interpolazione polinomiale, 34

Massimi e minimi, 66

Metodo dei minimi quadrati, 48

moltiplicatore, 15

norma matriciale, 5

Norme, 3

numero di condizionamento, 18

Ordine di convergenza, 62

ordine di convergenza, 62

Preliminari, 2

Proposizioni varie, 1

quadratura composita, 77

quadraturagaussiana, 77

raggio spettrale, 6

resto, 75

Ricerca di punti fissi, 57

riduzione di Gauss

naive, 14

Romberg, 79

Sistemi lineari, 13

Sistemi non lineari, 65

Spline, 45