Dispense di

Matematica Computazionale

Relative all'anno accademico 2009 – 2010

Corsi di laurea in:
Informatica
Tecnologie Informatiche

Docente: Prof.ssa ***B. Della Vecchia***

Rielaborazione e appunti a cura degli studenti:

Ciano Debora (debora.ciano87@alice.it)
D'Epifanio Stefano (sdepifanio@hotmail.it)
Gobbi Matteo (matteogobbi@matteogobbi.it)
Hima Erin (alfasite777@hotmail.com)
Perfetto Stefano (stefano.perfo@gmail.com)
Veschini Andrea (a.veschini@gmail.com)

Questi appunti sono soggetti ad aggiornamenti.
Per questo motivo gli studenti sono invitati a controllare periodicamente la pagina internet del corso in cerca di eventuali aggiornamenti.

Capitolo 1:
Sistemi lineari

- Metodi per la risoluzione di sistemi lineari

    - Metodi diretti
        - Metodo di eliminazione di Gauss
        - Fattorizzazione LU (Lower Up)
        - Risoluzione del sistema tramite fattorizzazione LU
            - La formula PA = LU
            - Risoluzione
            - Metodo di Cholesky
        - Calcolo della matrice inversa data una matrice triangolare inferiore
        - Il costo computazionale della triangolazione

    - Metodi iterativi
        - Metodo di Jacobi
        - Metodo di Gauss-Seidel
        - Metodo del sovrarilassamento (SOR)

Capitolo 2:
Autovalori

- Teorema di Gershgorin
- Metodo delle potenze
- Metodo delle potenze inverse

Capitolo 3:
Interpolazione polinomiale

- Motivazioni
- Che cos'è? Diamo una definizione matematica
- Prima soluzione: metodo dei coefficienti indeterminati
- Seconda soluzione: polinomi fondamentali di Lagrange
- Terza soluzione: polinomio interpolante di Lagrange espresso tramite rappresentazione di Newton (o metodo delle differenze divise)
- Stima dell'errore (resto dell'interpolazione)
- Convergenza dell'interpolazione
- Fenomeno di Runge

Capitolo 4:
Equazioni non lineari

- Metodi iterativi
    - Metodo di Newton Raphson
        - Interpretazione geometrica
        - Test di convergenza

Capitolo 4:
Quadratura numerica

- Formule di Newton-Cotes
    - Formula dei trapezi
    - Formula di Cavalieri-Simpson
    - Formule composite
- Formule di Gauss
    - Funzioni peso
- Strategie di quadratura numerica

Capitolo 5:
Approssimazione

- Introduzione
- Norma di Chebyshev
- Teorema di Weierstrass
- Norma quadratica
- I polinomi di Chebyshev
    - Zeri di Chebyshev
- Il problema della migliore approssimazione
    - Polinomio di quasi migliore approssimazione
    - Approssimazione ai minimi quadrati

Capitolo 6:
Malcondizionamento

- Norma
- Come vedere se un sistema è malcondizionato
- L'errore è sempre amplificato
- Esempio di problema condizionato

Capitolo 7:
Equazioni differeziali (del primo ordine)

- Introduzione
- Metodo di Eulero
    - Metodo di Eulero-Cauchy
- Metodi passo-passo (step by step)
    - Metodo di Milne
- Metodi di Runge - Kutta

Appendice:
Domande frequenti per l'esame

Conclusioni

Capitolo 1

1.1 SISTEMI LINEARI

Il problema fondamentale dell'algebra lineare è risolvere un *sistema di equazioni lineari*.
In matematica, e più precisamente in algebra lineare, un'*equazione lineare* è un'equazione di primo grado con un certo numero d'incognite. Un *sistema di equazioni lineari* (o sistema lineare) è un insieme di equazioni lineari, che devono essere verificate tutte contemporaneamente: in altre parole, una *soluzione* del sistema è tale se è soluzione di tutte le equazioni. La soluzione è quindi, l'insieme di valori $x_1 \ldots x_n$ che, sostituiti alle incognite, rende le equazioni delle *identità*.

In questo capitolo ci proponiamo di costruire metodi efficienti per la determinazione della soluzione di sistemi lineari di *n* equazioni in *n* incognite, cioè di sistemi lineari quadratici, essendo questo il caso più interessante e che si presenta nella maggior parte delle applicazioni, sia esplicitamente nel modello matematico associato al fenomeno fisico in esame, sia come passo intermedio o finale nella risoluzione numerica del modello in questione, rappresentato per esempio, da equazioni differenziali.

Cominciamo con un esempio:

$$\begin{cases} 2x - y = 0 \\ -x + 2y = 3 \end{cases}$$

A questo sistema di equazioni possiamo associare una *matrice* composta dai coefficienti delle incognite x e y. Una *matrice* non è altro che un *vettore rettangolare di numeri*.

Matrice dei coefficienti A

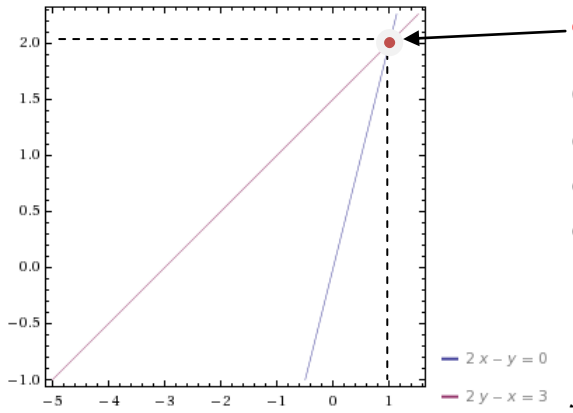
$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$$

b vettore dei termini noti

vettore delle incognite x

$$Ax = b \quad ; \quad A \in \mathbb{R}^{n \times n} \quad ; \quad b, x \in \mathbb{R}^n$$

Adesso ci proponiamo di risolvere il sistema lineare, cioè trovare x e y che soddisfano sia la prima sia la seconda equazione:



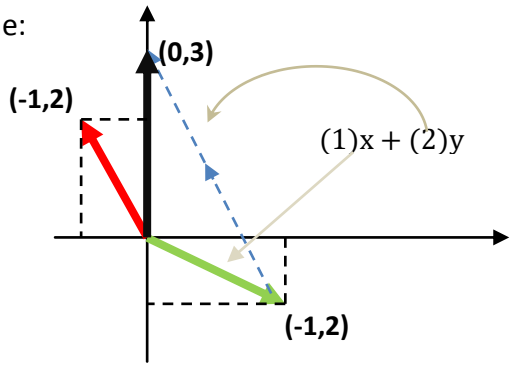
Quali punti soddisfano la prima equazione? Tutti quegli che giacciono sulla linea blu ovviamente (da qui equazione lineare). I punti che soddisfano la seconda equazione sono tutti e soli quegli che giacciono sulla linea rossa.

La soluzione del nostro sistema e il punto **c** ($x=1$, $y=2$) che giace su tutte e due le linee.

Interpretando invece la nostra matrice "*per colonne*" otteniamo un'altra "vista" del problema:

$$x\begin{pmatrix}2\\-1\end{pmatrix}+y\begin{pmatrix}-1\\2\end{pmatrix}=\begin{pmatrix}0\\3\end{pmatrix}$$

Ciò che questa equazione ci sta chiedendo e di trovare una giusta *combinazione lineare* tra il primo vettore $\begin{pmatrix}2\\-1\end{pmatrix}$ e il secondo $\begin{pmatrix}-1\\2\end{pmatrix}$ per ottenere il vettore $\begin{pmatrix}0\\3\end{pmatrix}$. Ed ecco la geometria che sta dietro l'equazione:



Ora ci facciamo una domanda. Siamo noi, in grado di risolvere questo sistema per ogni possibile valore del vettore di termine noto, cioè per ogni b? In questo caso la risposta e *si!* L'insieme di tutte le combinazioni lineari in questo caso finisce per ricoprire l'intero piano. Interpretando la matrice "per colonne" diventa chiaro che il sistema ammette soluzione se e solo se il vettore b appartiene allo spazio (lineare) generato dalle colonne di A.

1.2 METODI PER LA RISOLUZIONE DEI SISTEMI LINEARI

Di norma i metodi numerici per la risoluzione di sistemi lineari vengono suddivisi in due classi:

I. **Metodi diretti**
II. **Metodi Iterativi**

Con i *metodi diretti* l'esatta soluzione viene costruita in assenza di errori di arrotondamento in un numero finito di passi. Per sistemi con matrici A *dense*[1], i metodi diretti sono di solito i più efficienti.

I ***metodi iterativi*** invece sono generalmente utilizzati per la risoluzione di matrici A *sparse*[2], e di ordine elevato. Sistemi sparsi sono presenti in numerose applicazioni. A causa dell'elevato ordine delle matrici coinvolte in problemi di questo tipo (da alcune migliaia sino a 10^5 e oltre) i metodi diretti non sono sempre utilizzabili. Infatti, questi ultimi ottengono la soluzione x in un numero finito di passi mediante una successione (finita) di trasformazioni del problema iniziale in problemi equivalenti, cioè con la stessa soluzione x, ma con matrici dei coefficienti diverse; anzi, con il procedere del metodo di risoluzione, il numero di elementi non nulli presenti in queste matrici generalmente cresce, e può ben presto saturare lo spazio disponibile nella memoria centrale del calcolatore. In questi casi e utile, e spesso indispensabile, utilizzare metodi iterativi, i quali, operando sempre e solo con gli elementi della matrice iniziale A, generano una successione infinita di vettori convergente, sotto opportune condizioni, alla soluzione cercata. Poiché il processo iterativo lascia inalterata la matrice A, e sufficiente memorizzare gli elementi non nulli di A.

	Vantaggio	Svantaggio
I. Metodi diretti	-risoluzione del sistema in un numero finito di passi. -soluzione esatta.	-nel caso in cui ci trovassimo di fronte a matrici sparse(molti elementi =0) alterando la matrice ci ritroveremo con una matrice più complessa.
II. Metodi iterativi	-altera la matrice di partenza (quindi non è necessario salvarsi la matrice in memoria) -richiede una quantità costante di memoria	-soluzione approssimativa -costruzione di una successione di vettori convergenti.

[1] Una matrice densa è una matrice con pochi elementi pari a zero.
[2] Una matrice con molti elementi pari a zero è invece una matrice sparsa.

1.2.1 METODI DIRETTI

1. IL METODO DI ELIMINAZIONE DI GAUSS

Il metodo diretto più noto e più utilizzato è senza dubbio quello delle eliminazioni successive di Gauss.
Il metodo di Gauss è un metodo molto semplice e naturale per la risoluzione di un sistema lineare. E' il metodo che tutti (e anche i software package) usiamo ogni qual volta ci capita di risolvere un semplice sistema lineare.
Alla base di questo metodo vi è il presupposto secondo il quale risolvere una matrice che abbia forma triangolare superiore o inferiore sia molto semplice e immediato.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n = b_1 \\ a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots \\ a_{nn}x_n = b_n \end{cases}$$

Soluzione ultima riga della matrice:

$$x_n = \frac{b_n}{a_{nn}}$$

Soluzione della k-esima riga della matrice:

$$x_k = \frac{(b_k - \sum_{J=k+1}^{n} a_{kj}x_j)}{a_{kk}}$$

Questa considerazione ci suggerisce di esaminare la possibilità di trasformare un generico sistema non triangolare in un sistema equivalente(ovvero con le stesse soluzioni) di forma triangolare.

Se la matrice coinvolta è una matrice "buona" allora il metodo di eliminazione funzionerà e noi otterremmo la nostra risposta, ed anche in una maniera efficiente.
Vedremmo poi anche i casi, se esistono, che possono far fallire questo metodo.
Due sono i passi fondamentali che dobbiamo eseguire:

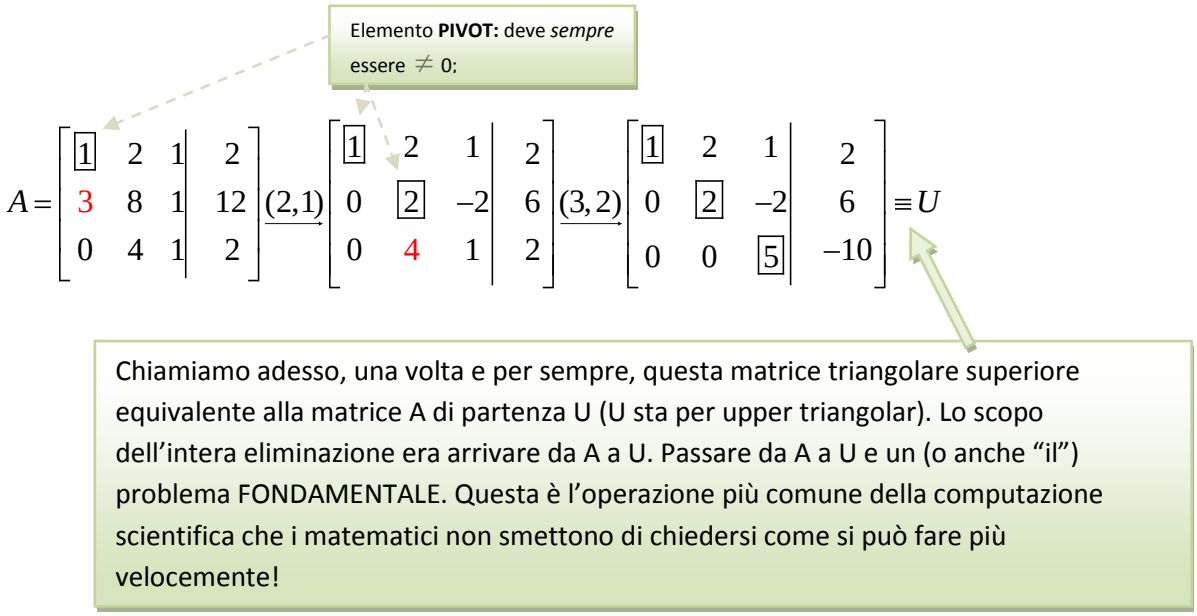
a) **Eliminazione**
b) **Sostituzione inversa(o permutazione)**

Cenni pratici e teorici sulle tecniche per operare su una matrice:

Facciamo un esempio teorico prima di procedere con la dimostrazione del metodo di Gauss per vedere come vengono applicati i due passi fondamentali citati in precedenza.. Questa volta ci proponiamo di risolvere il sistema lineare di 3 equazioni in 3 incognite seguente:

$$\begin{cases} x + 2y + z = 2 \\ 3x + 8y + z = 12 \\ 4y + z = 2 \end{cases}$$

a) Eliminazione



Chiamiamo adesso, una volta e per sempre, questa matrice triangolare superiore equivalente alla matrice A di partenza U (U sta per upper triangolar). Lo scopo dell'intera eliminazione era arrivare da A a U. Passare da A a U e un (o anche "il") problema FONDAMENTALE. Questa è l'operazione più comune della computazione scientifica che i matematici non smettono di chiedersi come si può fare più velocemente!

Il processo di eliminazione si porta avanti sostituendo successivamente le righe della matrice con la giusta combinazione lineare della riga in questione con le precedenti, dove per *"giusta combinazione lineare"* intendiamo quella che ci fa anulare l'elemento precedente(nella riga) al pivot.

Una combinazione lineare si ottiene ogni volta che:

1. Sostituiamo una riga con la stessa moltiplicata per un numero reale (scalare)
2. Sostituiamo una riga con la multipla di se stessa(operazione 1) + un'altra riga dello stesso sistema alla quale eventualmente gli abbiamo applicato l'operazione 1.

Formalmente:

Sia *V* uno spazio vettoriale su un campo *K*. Siano $v_1, \ldots, v_n$ vettori di *V*. Una **combinazione lineare** di questi è il vettore individuato dalla seguente scrittura

$$a_1 v_1 + a_2 v_2 + \ldots + a_n v_n$$

dove $a_1, \ldots, a_n$ sono scalari, cioè elementi di *K*. Gli scalari nella precedente espressione possono essere scelti ad arbitrio e sono detti **coefficienti** della combinazione lineare.

L'operazione 1 e 2 sonno gli unici operazioni legali che possiamo fare sulla matrice (operazioni che non cambiano la soluzione dell sistema).

b) Sostituzione inversa o premutazione

Se vogliamo permutare(o scambiare) due righe di una matrice identità basta moltiplicare quest'ultima con un'altra matrice per A: di fatto permuteremo le righe corrispondenti di A.

$$P_{1\leftrightarrow 2} \Leftarrow \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} * \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

$\Longleftarrow$ L'effetto della matrice $P_{1\text{-}2}$ è scambiare la prima riga con seconda.

Spiegazione del metodo di Gauss:

Prendiamo un generico sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots\ldots\ldots a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots\ldots\ldots a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots\ldots\ldots a_{3n}x_n = b_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases} \Longleftarrow (*)$$

e trasformiamolo in uno equivalente triangolare.

Supponiamo $a_{11} \neq 0$ (ipotesi che rilasceremo piu' avanti): possiamo eliminare l'incognita x_1 dalle ultime $(n-1)$ equazioni, cioe' dalla $2^a, 3^a, \ldots, n-esima$,sommando alla i-esima equazione , $i = 2,3,\ldots,n$ la prima moltiplicata per:

$$m_{i1} = -\frac{a_{i1}}{a_{11}} \quad , \quad i = 2,3,\ldots,n$$

Per esempio per i=2

$$-\frac{a_{21}}{a_{11}}(a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots\ldots\ldots a_{1n}x_n) = b_1(-\frac{a_{21}}{a_{11}})$$

$$-a_{21}x_1 - \frac{a_{21}}{a_{11}}a_{12}x_2 + \ldots\ldots\ldots = -b\frac{a_{21}}{a_{11}} \Longleftarrow (**)$$

Ora eseguo la somma tra la prima equazione del nostro sistema e la seconda che ho appena trovato(ovvero sommo (*) con (**)):

$$x_1(a_{21} - a_{21}) + x_2 a_{22}{}^{(2)} + x_3 a_{23} + \ldots\ldots = b_1{}^{(2)}$$

Riapplichiamo il procedimento (di eliminazione) alle ultime (n-1) equazioni. Se $a^{(2)}{}_{22} \neq 0$ possiamo eliminare l'incognita x_2 dalla $3^a, 4^a, \ldots n-esima$ equazione: e' sufficiente porre

$$m_{i2} = -\frac{a^{(2)}{}_{i2}}{a^{(2)}{}_{22}} \quad , \quad i = 3,\ldots,n$$

e sommare alla i-esima equazione la seconda moltiplicata per m_{i2}. Avremo un nuovo sistema equivalente a quello di partenza.

Dopo (n-1) passi arriveremo,supponendo tutti gli elementi di pivot non nulli, al seguente sistema triangolare:

$$\begin{cases} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + a_{13}^{(1)}x_3 + \cdots + a_{1n}^{(1)}x_n = b_1^{(1)} \\ a_{22}^{(2)}x_2 + a_{23}^{(2)}x_3 + \cdots + a_{2n}^{(2)}x_n = b_2^{(2)} \\ a_{33}^{(3)}x_3 + \cdots + a_{3n}^{(3)}x_n = b_3^{(3)} \\ \cdots\cdots\cdots\cdots\cdots \\ a_{nn}^{(n)}x_n = b_n^{(n)} \end{cases}$$

Lo scema di calcolo seguente riassume la descrizione del metodo di Gauss:

1) L'eliminazione delle variabili viene eseguita in (n-1) passi; al passo k-esimo, $k = 1, 2, \ldots, n-1$ gli elementi $a^{(k)}{}_{ij}$, i e $j > k$, e $b_i^{(k)}$ vengono trasformati in accordo con le formule

$$i = k+1, \ \ldots \ , n: \quad \begin{cases} m_{ik} = -\dfrac{a^{(k)}{}_{ik}}{a^{(k)}{}_{kk}} \\ a_{ij}{}^{(k+1)} = a_{ij}{}^{(k)} + m_{ik}a_{kj}{}^{(k)} \quad , \quad j = k+1, \ \ldots \ , n \\ b_i^{(k+1)} = bi^{(k)} + m_{ik}b_k{}^{(k)} \end{cases}$$

2) La soluzione del sistema triangolare finale risulta:

$$x_n = \frac{b_n}{a_{nn}}$$

$$x_k = \frac{(b_k - \sum_{J=k+1}^{n} a_{kj}x_j)}{a_{kk}}$$

$\Rightarrow$ **Possiamo risolvere il sistema con sole $\frac{n^2}{2}$ operazioni aritmetiche.**

N.B. Se l'elemento della prima riga della prima colonna $a_{1,1}$ fosse stato uguale a zero, allora mi sarei dovuto fermare e dire che non si può andare avanti? **No!**

In questo caso sarebbe possibile scambiare questa riga con un'altra del sistema e cosi facendo, avrei risolto il problema. Cioè se l'elemento al posto del pivot è zero, posso (e devo) scambiare(o permutare) l'equazione con un'equazione successiva del sistema.

Il metodo di eliminazioni di Gauss dimostra che e' sempre possibile mediante un numero finito di permutazioni e combinazioni lineari, trasformare un generico sistema, in un sistema triangolare[3].

Esempio di esecuzione del metodo di Gauss:

$$\begin{cases} 2x_1 - x_2 + x_3 - 2x_4 = 0 \\ 2x_2 - x_4 = 1 \\ x_1 - 2x_3 + x_4 = 0 \\ 2x_2 + x_3 + x_4 = 4 \end{cases} \Longrightarrow \begin{cases} 2x_1 - x_2 + x_3 - 2x_4 = 0 \\ 2x_2 - x_4 = 1 \\ \frac{1}{2}x_2 - \frac{5}{3}x_3 + 2x_4 = 0 \\ 2x_2 + x_3 + x_4 = 4 \end{cases}$$

$$\Longrightarrow \begin{cases} 2x_1 - x_2 + x_3 - 2x_4 = 0 \\ 2x_2 - x_4 = 1 \\ -\frac{5}{2}x_3 + \frac{9}{4}x_4 = -\frac{1}{4} \\ x_3 + 2x_4 = 3 \end{cases} \Longrightarrow \begin{cases} 2x_1 - x_2 + x_3 - 2x_4 = 0 \\ 2x_2 - x_4 = 1 \\ -\frac{5}{3}x_3 + \frac{9}{4}x_4 = -\frac{1}{4} \\ \frac{29}{10}x_4 = \frac{29}{10} \end{cases}$$

Partendo dall'ultima equazione (del sistema triangolare finale) otteniamo

$$\begin{aligned} x_4 &= 1 \\ x_3 &= \left(-\frac{1}{4} - \frac{9}{4}x_4\right)\left(-\frac{2}{5}\right) = 1 \\ x_2 &= (1 + x_4)/2 = 1 \\ x_1 &= (x_2 - x_3 + 2x_4)/2 = 1 \end{aligned}$$

Osservazioni finali

Il procedimento delle eliminazioni di Gauss può essere portato a termine senza permutare l'ordine iniziale delle equazioni, ovvero i successivi elementi pivot $a^{(k)}{}_{kk}$ sono tutti non nulli, se e solo se $\det(A_k) \neq 0, \quad k = 1, 2, \ldots, n$. Ciò e' senz'altro vero, per esempio, quando la matrice A e' a diagonale dominante per righe o per colonne($|a_{ii}| > \sum |a_{i,j}|$)oppure e' simmetrica definita positiva[4].

In alcuni casi però anche se il nostro pivot $a_{kk}^{(k)} \neq 0$ ma assume un valore molto piccolo è necessario, per **assicurare una maggiore stabilità numerica** permutare l'ordine delle equazioni

[3] Equivalente, cioe', con le stesse soluzioni del sistema originale.

[4] Una matrice è definita simmetrica e positiva quando è una matrice quadrata che ha la proprietà di essere la trsposta di se stessa e i suoi elementi sono tutti maggiori di zero.

usando la strategia del **pivot parziale:**ovvero dato un pivot generico a_{rk} scelgo r in modo che sia uguale al più piccolo intero ($a_{rk} = \max_{k \le i \le n} \left| a_{ik}^{(k)} \right|$) e successivamente scambierò l'equazione k-esima con la r-esima.

Se al generico passo k-esimo il processo di eliminazione non viene effettuato solo sulle righe successive alla k-esima, ma anche sulle precedenti, allora dopo n-passi otteniamo un sistema diagonale:

$$\begin{cases} a^{(1)}{}_{11}x_1 & 0 & \cdots & 0 & = \bar{b}_1^{(1)} \\ 0 & a^{(2)}{}_{22}x^2 & & \vdots & = \bar{b}_2^{(2)} \\ \vdots & \ddots & a^{(3)}{}_{33}x_3 & & = \bar{b}_3^{(3)} \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & 0 & a^{(n)}{}_{nn}x_n & = \bar{b}_n^{(n)} \end{cases}$$

Questa variazione del metodo di Gauss e' nota con il nome di ***Metodo Di Jordan***.

Il numero delle operazioni richieste dal Metodo di Gauss è $= \frac{n^3}{3}$

Il numero delle operazioni richieste dal Metodo Di Jordan è $= \frac{n^3}{2}$

3.Fattorizzazione Lower Up (LU)

Ci proponiamo di interpretare il metodo di Gauss come successione finita di trasformazione della matrice A e del termine noto b, cioè come moltiplicazioni di A e b per un numero finito di *opportune matrici.*
Questa interpretazione ci consentirà di riformulare l'algoritmo di Gauss in due parti distinte:

- la prima , la più costosa in termini di operazioni aritmetiche, ci determinerà una matrice non singolare G tale che $GA=U$, con U di forma triangolare superiore
- la seconda , utilizzando la matrice G, ci consentirà di risolvere il sistema Ax=b:

$$GAx = Gb \rightarrow Ux = \bar{b}$$

Osserviamo preliminarmente che lo scambio di due equazioni del sistema Ax=b (per esempio la i-esima con la j-esima), può essere interpretata come prodotto(da sinistra) di entrambi i membri del sistema per la matrice

La matrice Pi,j è una matrice di permutazione dove sono state scambiate le righe i e j.

$$P_{i,j} = \begin{bmatrix} 1 & 0 & 0 & .... & ... & 0 \\ 0 & 1 & 0 & .... & ... & 0 \\ 0 & 0 & 0 & 1 & ..... & 0 \\ 0 & 0 & 1 & 0 & ..... & 0 \\ 0 & 0 & 0 & 0 & 1 & ..... \end{bmatrix}$$

Riga j

Riga i

Analogamente, la sostituzione dell'equazione i-esima con la riduzione per la j-esima moltiplicata per il coefficiente $m_{i,j}$ può essere ottenuta moltiplicando Ax=b per la matrice

$$M_{i,j} = \begin{bmatrix} 1 & 0 & 0 & .... & ... & 0 \\ 0 & 1 & 0 & .... & ... & 0 \\ 0 & 0 & 1 & 0 & ..... & 0 \\ 0 & 0 & 0 & 1 & ..... & 0 \\ 0 & 0 & 0 & 0 & 1 & ..... \end{bmatrix}$$

Quindi con il metodo di Gauss determino implicitamente delle matrici $P_1, P_2, ...P_{n-1}$ di tipo I(identità) quando non avvengono scambi di equazioni e $P_{i,j}$ (matrici di permutazioni)altrimenti, e delle matrici $M_1, M_2, ...M_{n-1}$, con

$$M_j = M_{nj} .....M_{j+2,j} M_{j+1,j}$$

tale che il nuovo sistema $M_{n-1}P_{n-1} - M_i P_i Ax = M_{n-1}P_{n-1} .....M_1 P_1 b$

assuma la forma triangolare superiore $Ux = \bar{b}$, ossia

$$M_{n-1}P_{n-1} .....M_1 P_1 A = U$$

N.B. Posto $G = M_{n-1}P_{n-1} .....M_1 P_1$ denotiamo questa decomposizione con $GA = U$

N.B.2 Ricordiamo che questa decomposizione ha come costo computazionale $\frac{n^3}{3}$

Osservazioni

(i) Nelle applicazioni è superfluo costruire esplicitamente G poiché G verrà utilizzata solo per trasformare matrici e vettori: è quindi sufficiente memorizzarsi i coefficienti $m_{i,j}$ e le permutazioni effettuate. I molteplici $m_{i,j}$ verranno memorizzati nelle corrispondenti posizioni della matrice A e al termine della triangolarizzazione al posto della matrice iniziale A avremo

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ m_{21} & u_{22} & u_{23} & \cdots & u_{2n} \\ m_{31} & m_{32} & u_{33} & \cdots & u_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ m_{n1} & m_{n2} & m_{n3} & \cdots & u_{nn} \end{pmatrix}$$

(ii) Anche la costruzione di $P_{i,j}$ è superflua. Poiché la loro unica funzione è di provocare scambi di righe, possiamo riprodurre tali azioni semplicemente memorizzando gli scambi effettuati in un altro valore pivot di n-1 componenti.
Per esempio se al passo k viene effettuata la permutazione tra la riga k e la riga r pongo *pivot(k)=r*, invece se la riga k-esima rimane inalterata allora avrò *pivot(k)=k*.

In definitiva: Note le matrici G e U, per risolvere il sistema Ax=b è sufficiente porre $GAx=Gb$ cioè

$$\bar{b} = Gb = M_{n-1}P_{n-1}.....M_1P_1b$$

4.RISOLUZIONE DEL SISTEMA TRAMITE LA FATTORIZZAZIONE LU

4.1 La formula PA=LU

La riformulazione del problema di Gauss in due fasi divise(nella prima trasformiamo solo la matrice A, mentre nella seconda costruiamo il vettore $\bar{b} = Gb$ e risolviamo il sistema triangolare $Ux = \bar{b}$)ci consente di risolvere in modo efficiente sistemi del tipo

$$\begin{cases} Ax_1 = b_1 \\ Ax_2 = b_2 \\ \vdots \\ \vdots \\ Ax_n = b_n \end{cases}$$

ATTENZIONE:VANTAGGIO !
E' la stessa matrice che và decomposta una sola volta

Supponiamo di aver già determinato la decomposizione di Gauss. La conoscenza delle matrici M_i, P_i ci consente di riordinare le successive trasformazioni di A:

Questo è proprio G

$$M_{n-1}P_{n-1}.....M_1P_1A = (\overline{M}_{n-1}....\overline{M}_2\overline{M}_1)(P_{n-1}.....P_2P_1)A$$

così che posto $\overline{M} = \overline{M}_{n-1}....\overline{M}_2\overline{M}_1$ e $P = P_{n-1}.....P_2P_1$

possiamo scrivere:

$$\overline{M}PA = U$$

$$PA = \overline{M}^{-1}U$$

Di conseguenza possiamo affermare pertanto che il metodo di Gauss può essere utilizzato anche per determinare una matrice di permutazione P, una matrice triangolare inferiore con diagonale unitaria L(chiamato $L = \overline{M}^{-1}$) e una matrice triangolare superiore U, tali che

$$PAx = Pb$$
$$LUx = Pb$$

4.2 Risoluzione

Nota la fattorizzazione LU per determinare la soluzione del sistema $Ax=b$ è sufficiente risolvere i seguenti 2 sistemi triangolari

$$PAx = Pb \qquad Ly = Pb$$
$$LUx = Pb \qquad Ux = y$$

Se il metodo di Gauss non richiede scambio di righe sappiamo come $P=I$ e quindi $A = LU = LDU_1$ dove L e U_1 che sono matrici triangolari con diagonali unitarie e D è una matrice che contiene tutti gli elementi principali che apparivano in U.

Metodo di Cholesky

Quando A è simmetrica abbiamo $U_1 = L^T$ (cioè la trasposta di L), inoltre gli elementi $(D)_{ii}$ sono tutti positivi se e solo se A è definita positiva(ovvero non ci sono state permutazioni).Questo fatto ci consente di concludere che esiste allora un'unica matrice triangolare inferiore L_1, con elementi diagonali positivi, tale che $A = L_1 L_1^T$ dove $L_1 = LD^{1/2}$ dove con $D^{1/2}$ denotiamo la matrice diagonale che ha come elementi diagonale $(D^{1/2})_{ii} = \sqrt{(D)_{ii}}$.

Gli elementi di questa matrice possono venire determinati dalle seguenti formule:

$$L_1 = \begin{bmatrix} l_{11} & 0 & 0 & \dots & 0 \\ l_{21} & l_{22} & 0 & \dots & 0 \\ l_{31} & l_{32} & l_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ l_{n1} & l_{n2} & l_{n3} & \dots & l_{nn} \end{bmatrix}$$

$$l_{ii} = \sqrt{a_{ii}}$$

$$l_{i,j} = (a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}) / l_{jj}$$

$$l_{ii} = (a_{ii} - \sum_{k=1}^{i=1} l_{ik}^2)^{1/2}$$

Il costo computazionale di questo metodo è $\frac{n^3}{6}$ operazioni ovvero **proprio la metà** rispetto alla fattorizzazione LU

5. CALCOLO DELLA MATRICE INVERSA DATA UNA MATRICE TRIANGOLARE INFERIORE

Data la matrice triangolare inferiore

$$L = \begin{bmatrix} e_{11} & 0 & 0 & \dots & 0 \\ e_{21} & e_{22} & 0 & \dots & 0 \\ e_{31} & e_{32} & e_{33} & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ e_{n1} & e_{n2} & e_{n3} & \dots & e_{nn} \end{bmatrix}$$

Possiamo calcolare la sua matrice inversa

$$Y = L^{-1} = \begin{bmatrix} y_{11} & 0 & 0 & \dots & 0 \\ y_{21} & y_{22} & 0 & \dots & 0 \\ y_{31} & y_{32} & y_{33} & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ y_{n1} & y_{n2} & y_{n3} & \dots & y_{nn} \end{bmatrix}$$

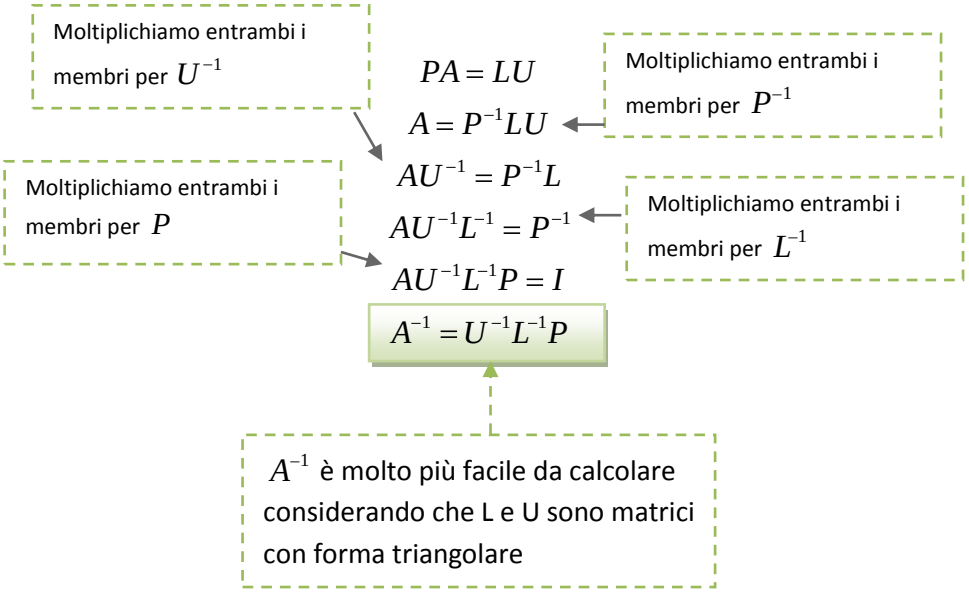
tramite questo semplice **algoritmo:**

$$y_{jj} = \frac{1}{e_{jj}} \qquad \textit{Per j =1.....n}$$

$$y_{ij} = \frac{\sum_{h=j}^{i-1} e_{ik} y_{kj}}{e_{ii}}$$

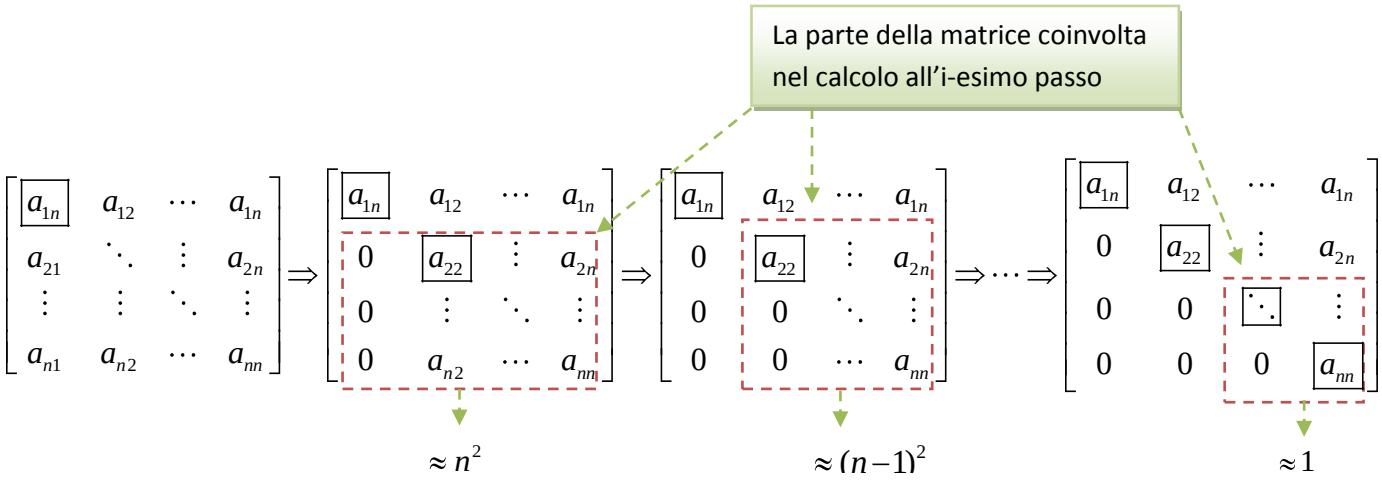
N.b.*Ma perché è cosi importante il calco della matrice inversa??*

Perché il calcolo dell'inversa della matrice A è legato al calcolo della matrice inversa di L e U



6.Il Costo Computazionale Della Triangolazione

Riflettiamo adesso sul costo del metodo di Gauss per passare da una matrice A qualunque a una matrice triangolare superiore equivalente U .



$\Rightarrow$ Il costo computazionale per passare da A a U e' $\approx n^2 + (n-1)^2 + \cdots + 1^2 \approx \frac{1}{3}n^3$

$\frac{1}{3}n^3$ e l'integrale di n^2 ($\int_1^n x^2 dx$).

Il costo della colonna dei termini noti che ci portiamo dietro durante l'eliminazione è n^2 .

OSSERVAZIONE

Abbiamo precisato che durante la fase dell'eliminazione, non accettiamo pivot nulli. Per la verità non solo non accettiamo pivot nulli, ma non acetiamo neanche pivot tropo piccoli o tropo grandi[5]. Tale esigenza discende dal fatto che pivot tropo piccoli/grandi potrebbero introdurre nel sistema instabilità numerica.(vedi osservazione del paragrafo sul metodo di Gauss).

1.2.3 METODI ITERATIVI

In alcune situazioni, per esempio nella soluzione di equazioni a derivate parziali di tipo ellittico con metodi alle differenze finite o agli elementi finiti, i sistemi da risolvere sono sparsi e di dimensioni tali ($(n \approx 10^3 \to 10^6)$ da rendere inutilizzabile, o quanto meno inefficiente, il metodo di Gauss anche con i moderni calcolatori di grande capacita. Infatti, mentre in questi casi la matrice iniziale ha un numero di elementi non nulli $p \ll n^2$, il processo delle eliminazioni successive del metodo di Gauss cambia le equazioni del sistema ad ogni passo, cosi che la matrice dei coefficienti può diventare sempre meno sparsa e richiedere quindi la memorizzazione di un numero eccessivo di elementi.
I metodi iterativi invece non alterano mai la matrice iniziale A.
Partendo da un'approssimazione iniziale $x^{(0)}$ essi definiscono una successione di approssimazioni $x^1, x^2, \ldots$ convergente, sotto opportune ipotesi, alla soluzione x del sistema non singolare[6]

$$Ax = b.$$

1. METODO DI JACOBI

Scriviamo esplicitamente il sistema $Ax = b$:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x^1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

con le equazioni ordinate in modo tale che $a_{ii} \neq 0 \quad i = 1, 2, \ldots, n$

proviamo quindi a risolvere il nostro sistema cercando i valori delle nostre incognite:

$$x_1 = \frac{b_1 - \sum_{j \neq 1}^{n} a_{1j}x_j}{a_{11}} \qquad x_2 = \frac{b_2 - \sum_{j \neq 2}^{n} a_{2j}x_j}{a_{22}}$$

[5] Questa è una necessita del calcolo numerico, non dell'algebra.

[6] O meglio, alla soluzione $\bar{x}$ del sistema perturbato $\bar{A}\bar{x} = \bar{b}$

$$x_3 = \frac{b_3 - \sum_{j\neq 3}^{n} a_{3j}x_j}{a_{33}} \qquad\qquad x_i = \frac{b_i - \sum_{j\neq i}^{n} a_{ij}x_j}{a_{ii}}$$

$$x^{(0)} = (x_1^{(0)}, x_2^{(0)}, ..., x_n^{(0)})$$

Come possiamo dedurre dall' esempio la mia incognita x_1 è legata ad altre x.

Allora applichiamo il **metodo di Jacobi** il quale consiste nel calcolare, nota un'approssimazione iniziale $x^{(0)}$ (oppure partendo da $x^{(0)} = 0$[7]), le approssimazioni successive:

$$x^{(k+1)} = (x_1^{(k+1)}, x_2^{(k+1)}, \ldots, x_n^{(k+1)}) \qquad k = 0,\ 1,\ 2, \ldots$$

inseriamo i valori del vettore iniziale nelle formule:

$$x_1^{(1)} = \frac{b_1 - \sum_{j\neq 1}^{n} a_{1j}x_j^{(0)}}{a_{11}} \qquad x_2^{(1)} = \frac{b_2 - \sum_{j\neq 2}^{n} a_{2j}x_j^{(0)}}{a_{22}} \qquad x_3^{(1)} = \frac{b_3 - \sum_{j\neq 3}^{n} a_{3j}x_2^{(0)}}{a_{33}}$$

con il nuovo vettore delle mie approssimazioni ed i nuovi valori eseguo lo stesso processo:

$$x_1^{(2)} = \frac{b_1 - \sum_{j\neq 1}^{n} a_{1j}x_j^{(1)}}{a_{11}} \qquad x_2^{(2)} = \frac{b_2 - \sum_{j\neq 2}^{n} a_{2j}x_j^{(1)}}{a_{22}} \qquad x_3^{(2)} = \frac{b_3 - \sum_{j\neq 3}^{n} a_{3j}x_j^{(1)}}{a_{33}}$$

Come è possibile vedere il metodo allora si può sintetizzare in questa **relazione**:

$$x_i^{(k+1)} = \frac{b_i - \sum_{j=1, j\neq i}^{n} a_{ij}x_j^{(k)}}{a_{ii}} \qquad i = 1,\ 2, \ldots, n$$

N.b. L'ipotesi seconda la quale questo metodo converga è che la nostra matrice A sia a diagonale dominante(ovvero quando $|a_{ii}| > \sum |a_{i,j}|$)

[7] Questo non è uno zero ma bensì indica come il nostro vettore delle approssimazioni iniziale sia il **vettore nullo**

2. METODO DI GAUSS-SEIDEL

Nel metodo di Jacobi ogni singola componente di $x^{(k+1)}$ dipende unicamente dall'approssimazione precedente $x^{(k)}$.

$$x_1^{(2)} = \frac{b_1 - \sum_{j\neq 1}^{n} a_{1j} x_j^{(1)}}{a_{11}} \qquad x_2^{(2)} = \frac{b_2 - \sum_{j\neq 2}^{n} a_{21} x_1^{(1)}}{a_{22}} \qquad x_3^{(2)} = \frac{b_3 - \sum_{j\neq 3}^{n} a_{32} x_2^{(1)}}{a_{33}}$$

Possiamo notare però come quando andiamo a calcolare $x_2^{(2)}$ io inserisca nella $x_j^{(1)}$ il valore di $x_1^{(1)}$ anche se abbiamo calcolato già il valore di $x_1^{(2)}$. La stessa cosa vale per $x_3^{(2)}$. Allora più genericamente possiamo dire che calcolata $x^{(k+1)}$ potremmo già utilizzare questo nuovo valore nella determinazione di $x_2^{(k+1)}$,e poi utilizzare $x_1^{(k+1)}$ e $x_2^{(k+1)}$ (più $x_4^{(k)},\ldots,x_n^{(k)}$) nel calcolo di $x_3^{(k+1)}$, e cosi via.

Questo procedimento rappresenta il **metodo di Gauss-Seidel.**

$$x_i^{(k+1)} = \frac{b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)}}{a_{ii}} \qquad i = 1,\ 2,\ldots,n$$

Anche questo metodo, come quello di Jacobi, converge sia quando la nostra matrice è a diagonale dominante per righe ($|a_{ii}| > \sum |a_{i,j}|$)o a diagonale dominante per colonne($|a_{kk}| > \sum |a_{ik}|$)

3. METODO DI SOVRARILASSAMENTO (SOR)

Riprendiamo il metodo di Gauss-Seidel. Dalla relazione (*) sottraendo $x_i^{(k)}$ da ambo i membri, otteniamo:

$$r_i^{(k)} = x_i^{(k+1)} - x_i^{(k)} = \frac{1}{a_{ii}}\left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij} x_j^{(k)} \right] \qquad i = 1,\ 2,\ldots,n$$

dove $r_i^{(k)}$ rappresenta la correzione da apportare a $x_i^{(k)}$ per ottenere la nuova approssimazione:

$$x_i^{(k+1)} = x_i^{(k)} + \omega r_i^{(k)} \qquad k = 0,\ 1,\ 2,\ldots$$

La formula scritta sopra definisce un **nuovo metodo** detto di **rilassamento**. Scegliere il parametro ω non è cosi semplice in quanto bisognerà sceglierlo in modo da accelerare il più possibile la convergenza della nostra successione $x^{(k)}$. Per esempio se $1 < \omega \leq 2$ il valore ottenuto dall'iterazione attuale viene incrementato per il valore assoluto: si sta assumendo che la soluzione si stia muovendo **troppo lentamente verso la convergenza**, e in questo caso si parla di

sovrarilassamento. Da mettere in risalto è il caso particolare in cui ω **sia uguale a 1: in questo caso il metodo si identifica con quello di Gauss_Seidel**

Dimostrazione:

$$x_i^{(k+1)} = x_i^{(k)} + 1 \cdot r_i^{(k)}$$

$$x_i^{(k+1)} = x_i^{(k)} + r_i^{(k)}$$

ma: $$r_i^{(k)} = x_i^{(k+1)} - x_i^{(k)} = \frac{1}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=1}^{n} a_{ij} x_j^{(k)} \right] \qquad i = 1,\ 2, \ldots, n$$

quindi: $$x_i^{(k+1)} = \frac{b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)}}{a_{ii}}$$

Cioè proprio la formula di Gauss- Siedel

Rewrite by Matteo Gobbi – Matricola: 1358489
Per esame di Matematica Computazionale
tenuto dalla Dott.ssa Biancamaria Della Vecchia

Capitolo 2

Autovalori

Gli autovalori, e di conseguenza gli autovettori, possono esserci molto utili in matematica. Ad esempio per diagonalizzare una matrice, oppure per poter scoprire se c'è qualcosa che rimane immutato dopo l'applicazione di una trasformazione lineare, etc.

Dati:

- $A \in \mathbb{R}^{n*n}$ matrice *n* righe ed *n* colonne
- $\lambda \in \mathbb{C}$ scalare

se:

- *(2.0.1)* $Ax = \lambda x, x \neq 0$ dove ***x*** è un vettore non nullo

allora λ si dice ***AUTOVALORE delle matrice A*** e ***x il corrispondente AUTOVETTORE***.

L'equazione *(2.0.1)* , può essere riscritta come:

$$Ax - \lambda x = 0, x \neq 0$$

e successivamente, raccogliendo *x* come:

(2.0.2) $$(A - \lambda I)x = 0, x \neq 0$$

Pertanto richiediamo che il sistema *(2.0.2)* ammetta soluzioni diverse dalla soluzione nulla , cioè il sistema *(2.0.2)* ammetta più di una soluzione e questo dall'algebra sappiamo che equivale a richiedere che:

(2.0.3) $$det(A - \lambda I) = 0$$

(Se questo determinante fosse diverso da zero, allora il sistema *(2.0.2)* ammetterebbe solo una soluzione, quella banale *x=0*, che noi non vogliamo).
Sviluppando dunque il determinante *(2.0.3)* otteniamo:

$$det(A - \lambda I) = \lambda^n + \alpha_1\lambda^{n-1} + \alpha_2\lambda^{n-2} + ... + \alpha_{n-1}\lambda + \alpha_n$$

Pertanto gli autovalori λ_i coincidono con le n radici dell'equazione caratteristica *(2.0.3)*.

Questa relazione sembra suggerire la determinazione degli autovalori, come le radici **dell'equazione caratteristica** .Tuttavia questo approccio risulta poco efficiente, pertanto sono stati sviluppati metodi ad hoc.
Per scegliere un metodo efficiente, occorre rispondere alle seguenti domande:

- è richiesto solo l'autovalore più grande in modulo ed il corrispondente autovettore?
- sono richiesti tutti gli autovettori e gli autovalori corrispondenti?
- la matrice ha proprietà speciali (simmetrica, tridiagonale, sparsa) ?

A seconda delle risposte a queste domande possiamo individuare metodi più efficienti di altri.

In una matrice $A \in \mathbb{R}^{n*n}$ avremo dunque *n* **autovalori** ed *n* **autovettori**, dato che**:**

$$\begin{aligned} Ax_1 &= \lambda_1 x_1 \\ Ax_2 &= \lambda_2 x_2 \\ &..... \\ &..... \\ &..... \\ Ax_n &= \lambda_n x_n \end{aligned}$$

dove appunto:

$x_1, x_2, ..., x_n$ sono gli **AUTOVETTORI**

$\lambda_1, \lambda_2, ..., \lambda_n$ sono i corrispondenti **AUTOVALORI**

2.1 Teorema di Gershgorin

Diamo ora un criterio semplice per la localizzazione degli autovalori di una matrice A. Definiamo:

- somma dei valori assoluti degli elementi fuori della diagonale, della riga *i*-esima:

$$r_i = \sum_{j \neq i, j=1}^{n} |a_{ij}|, i = 1, 2, ..., n$$

- somma dei valori assoluti degli elementi fuori della diagonale, della colonna *j*-esima:

$$c_j = \sum_{i \neq j, i=1}^{n} |a_{ij}|, j = 1, 2, ..., n$$

Chiamiamo dunque, **Cerchi di Gershgorin**, i cerchi del piano complesso di centro a_{ii} e raggio r_i:

$$R_i = \{z : |z - a_{ii}| \leq r_i, z \in \mathbb{C}\}$$

e i cerchi del piano complesso di centro a_{jj} e raggio c_j :

$$C_i = \{z : |z - a_{jj}| \leq c_j, z \in \mathbb{C}\}$$

Infine, definiamo i due insiemi *R* e *C*, rispettivamente unione di tutti i cerchi R_i e C_j:

$$R = \bigcup_{i=1}^{n} R_i \qquad C = \bigcup_{j=1}^{n} C_j$$

Siamo ora pronti per enunciare il teorema di Gershgorin:

Teorema: *Data una matrice A e gli insiemi R e C (appena definiti), allora gli autovalori di A appartengono all'insieme* $R \cap C$.

Ogni componente di *R* o di *C,* cioè ogni unione connessa massimale di cerchi *Ri* o *Cj*, contiene tanti autovalori di *A,* quanti sono i cerchi della componente, tenendo conto della molteplicità di ogni autovalore e di ogni cerchio.

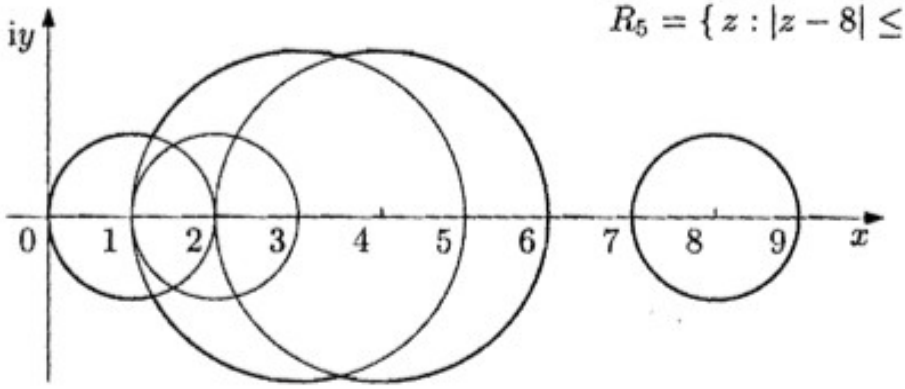
<u>Vediamo ora un esempio:</u>

Applichiamo il teorema alla matrice:

$$A = \begin{pmatrix} 4 & -1 & 1 & 0 & 0 \\ 1 & 3 & -1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 8 \end{pmatrix}$$

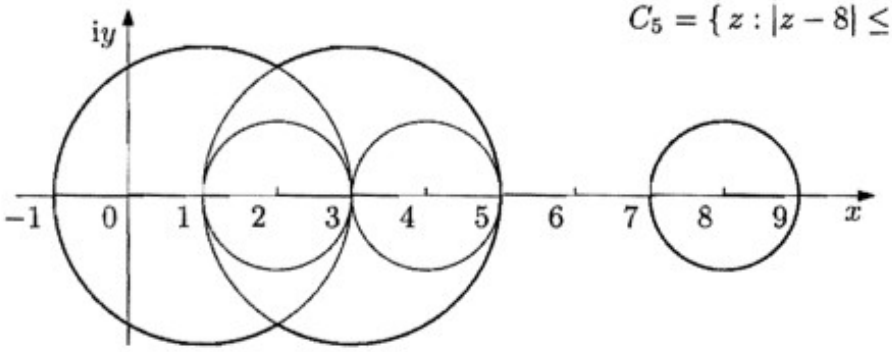
i cui autovalori sono $\lambda_1 = 5 + \sqrt{10},\ \lambda_2 = \lambda_3 = 3,\ \lambda_4 = 2,\ \lambda_5 = 5 - \sqrt{10}.$
Tutti gli autovalori appartengono all'insieme *R* unione dei cerchi seguenti:

$$R_1 = \{ z : |z - 4| \leq 2 \}$$
$$R_2 = \{ z : |z - 3| \leq 2 \}$$
$$R_3 = \{ z : |z - 1| \leq 1 \}$$
$$R_4 = \{ z : |z - 2| \leq 1 \}$$
$$R_5 = \{ z : |z - 8| \leq 1 \}$$

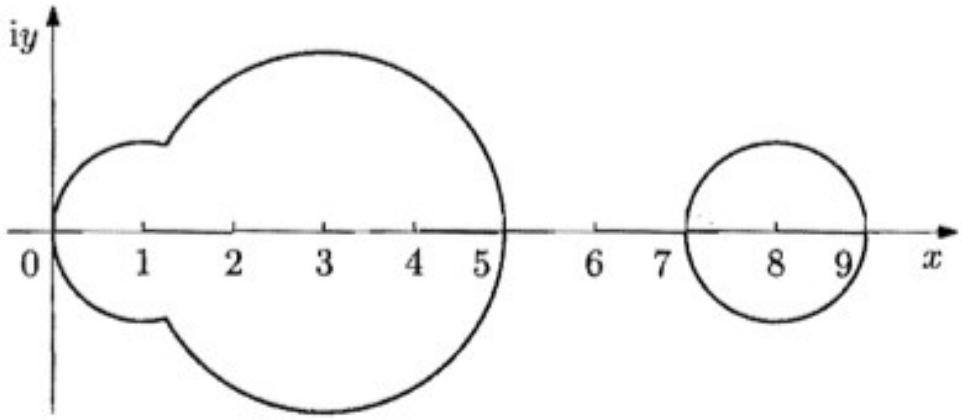


La regione *R* è formata dalle due componenti segnate in neretto. In una vi sono 4 cerchi e quindi 4 autovalori; l'altra, essendo formata da 1 solo cerchio, contiene 1 autovalore. Tutti gli autovalori appartengono però anche alla regione *C* unione dei cerchi *Cr*:

$$C_1 = \{ z : |z-4| \leq 1 \}$$
$$C_2 = \{ z : |z-3| \leq 2 \}$$
$$C_3 = \{ z : |z-1| \leq 2 \}$$
$$C_4 = \{ z : |z-2| \leq 1 \}$$
$$C_5 = \{ z : |z-8| \leq 1 \}$$



La regione *C* (segnata in neretto) ha due componenti: la prima contiene 4 autovalori, la seconda 1.
Possiamo pertanto concludere che tutti gli autovalori devono giacere nella regione intersezione $R \cap C$.



2.2 Metodo delle Potenze

Il *metodo delle potenze*, è un metodo utilizzato per determinare l'autovalore di modulo massimo. Questo autovalore gioca un ruolo cruciale in molti problemi numerici, ad esempio nel calcolo del numero di condizionamento di un sistema lineare. Tale metodo è applicabile sotto l'ipotesi in cui:

1. esista **un solo** autovalore di modulo massimo:

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

2. $x_1, x_2, ..., x_n$ siano autovettori linearmente indipendenti.

Sotto queste ipotesi siamo certi che:

$$\lambda_1 \in \mathbb{R}, \lambda_1 \notin \mathbb{C}$$

cioè che il nostro autovalore di modulo massimo, sia un numero reale non complesso, poiché se lo fosse, allora in modulo sarebbe uguale al suo coniugato, **contraddicendo** dunque **l'unicità**, **ovvero le nostre ipotesi**:

$$|x + yi| = |x - yi| = x$$

Andiamo dunque ad applicare il metodo:

1. Prendiamo un vettore v_0 qualsiasi, e scriviamolo come combinazione lineare degli n autovettori $x_1, x_2, ..., x_n$ (possiamo farlo per l'ipotesi 2):

$$v_0 = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n$$

(possiamo sempre supporre α_1 diverso da zero).

2. Scriviamo il vettore v_1 come il prodotto della matrice *A* per il vettore v_0 :

$$v_1 = Av_0 = A(\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n) = \alpha_1 Ax_1 + \alpha_2 Ax_2 + ... + \alpha_n Ax_n$$

3. Ricordiamo che la definizione di autovalore per il generico λ_i vale:

$$Ax_i = \lambda_i x_i$$

e quindi sostituendo il valore Ax_i con $\lambda_i x_i$, otteniamo:

$$v_1 = \alpha_1\lambda_1 x_1 + \alpha_2\lambda_2 x_2 + ... + \alpha_n\lambda_n x_n$$
$$= \lambda_1(\alpha_1 x_1 + \alpha_2\frac{\lambda_2}{\lambda_1}x_2 + ... + \alpha_n\frac{\lambda_n}{\lambda_1}x_n)$$

4. Ripetiamo il passo 2, ma questa volta per il vettore v_2 :

$$v_2 = Av_1 = A(\alpha_1\lambda_1 x_1 + \alpha_2\lambda_2 x_2 + ... + \alpha_n\lambda_n x_n) =$$
$$\alpha_1\lambda_1 Ax_1 + \alpha_2\lambda_2 Ax_2 + ... + \alpha_n\lambda_n Ax_n$$
$$= \lambda_1(\alpha_1 Ax_1 + \alpha_2\frac{\lambda_2}{\lambda_1}Ax_2 + ... + \alpha_n\frac{\lambda_n}{\lambda_1}Ax_n)$$

5. Quindi sempre in vigore del punto 3, sostituendo il valore Ax_i con $\lambda_i x_i$, otteniamo:

$$v_2 = \alpha_1\lambda_1^2 x_1 + \alpha_2\lambda_2^2 x_2 + ... + \alpha_n\lambda_n^2 x_n$$
$$= \lambda_1^2(\alpha_1 x_1 + \alpha_2\frac{\lambda_2^2}{\lambda_1^2}x_2 + ... + \alpha_n\frac{\lambda_n^2}{\lambda_1^2}x_n)$$

6. E ancora:

$$v_3 = \lambda_1^3(\alpha_1 x_1 + \alpha_2\frac{\lambda_2^3}{\lambda_1^3}x_2 + ... + \alpha_n\frac{\lambda_n^3}{\lambda_1^3}x_n)$$

7. Ripetiamo lo stesso procedimento. Otteniamo quindi l'espressione del vettore *k*-esimo:

$$v_k = \alpha_1\lambda_1^k x_1 + \alpha_2\lambda_2^k x_2 + ... + \alpha_n\lambda_n^k x_n$$
$$= \lambda_1^k(\alpha_1 x_1 + \alpha_2\frac{\lambda_2^k}{\lambda_1^k}x_2 + ... + \alpha_n\frac{\lambda_n^k}{\lambda_1^k}x_n)$$

8. Non essendo definita l'operazione di rapporto fra vettori, procediamo calcolando il rapporto delle coordinate *j*-esime dei vettori v_k e $v_{(k+1)}$, ovvero:

(2.2.1) $$\frac{(v_{k+1})_j}{(v_k)_j} = \frac{\lambda_1^{k+1}(\alpha_1 x_1 + \alpha_2 \frac{\lambda_2^{k+1}}{\lambda_1^{k+1}} x_2 + \ldots + \alpha_n \frac{\lambda_n^{k+1}}{\lambda_1^{k+1}} x_n)_j}{\lambda_1^{k}(\alpha_1 x_1 + \alpha_2 \frac{\lambda_2^{k}}{\lambda_1^{k}} x_2 + \ldots + \alpha_n \frac{\lambda_n^{k}}{\lambda_1^{k}} x_n)_j}$$

9. Calcoliamo il $\lim_{k\to+\infty}$ del rapporto al primo membro in *(2.2.1)*:

$$\lim_{k\to+\infty} \frac{(v_{k+1})_j}{(v_k)_j} = \lim_{k\to+\infty} \frac{\lambda_1^{k+1}(\alpha_1 x_1 + \alpha_2 \frac{\lambda_2^{k+1}}{\lambda_1^{k+1}} x_2 + \ldots + \alpha_n \frac{\lambda_n^{k+1}}{\lambda_1^{k+1}} x_n)_j}{\lambda_1^{k}(\alpha_1 x_1 + \alpha_2 \frac{\lambda_2^{k}}{\lambda_1^{k}} x_2 + \ldots + \alpha_n \frac{\lambda_n^{k}}{\lambda_1^{k}} x_n)_j}$$

$$\frac{\lim_{k\to+\infty}(v_{k+1})_j}{\lim_{k\to+\infty}(v_k)_j} = \frac{\lim_{k\to+\infty} \lambda_1^{k+1}(\alpha_1 x_1 + \alpha_2 \frac{\lambda_2^{k+1}}{\lambda_1^{k+1}} x_2 + \ldots + \alpha_n \frac{\lambda_n^{k+1}}{\lambda_1^{k+1}} x_n)_j}{\lim_{k\to+\infty} \lambda_1^{k}(\alpha_1 x_1 + \alpha_2 \frac{\lambda_2^{k}}{\lambda_1^{k}} x_2 + \ldots + \alpha_n \frac{\lambda_n^{k}}{\lambda_1^{k}} x_n)_j}$$

10. Osservando che:

$$\frac{\lambda_i^{k+1}}{\lambda_1^{k+1}} < 1, i \neq 1$$

segue che se $k \to +\infty$, risulta quindi:

$$\lim_{k\to+\infty} \frac{\lambda_i^{k+1}}{\lambda_1^{k+1}} = 0$$

per cui:

$$\frac{\lim_{k\to+\infty}(v_{k+1})_j}{\lim_{k\to+\infty}(v_k)_j} = \frac{\lim_{k\to+\infty} \lambda_1^{k+1}(\alpha_1 x_1)_j}{\lim_{k\to+\infty} \lambda_1^{k}(\alpha_1 x_1)_j} = \lambda_1$$

La convergenza della successione:

$$\frac{(v_{k+1})_j}{(v_k)_j}$$

al limite λ_1 , dipende dalla potenza:

$$(\frac{\lambda_1}{\lambda_2})^k$$

La convergenza è tanto più rapida quanto più piccolo è il rapporto:

$$\left|\frac{\lambda_1}{\lambda_2}\right|$$

Abbiamo quindi trovato un modo per calcolare l'autovalore di modulo massimo.

2.3 Metodo delle potenze inverse

Ci proponiamo di raffinare un'approssimazione p di un autovalore λ (cioè conosciamo un valore approssimato grossolanamente p , per l'autovalore λ e voglio avere un'approssimazione più accurata). Partiamo dalla definizione di autovalore:

$$Ax = \lambda x$$

e moltiplichiamo ambo i membri dell'equazione per A^{-1} :

$$A^{-1}Ax = A^{-1}\lambda x$$

la quale diventa quindi:

$$x = A^{-1}\lambda x$$

Ora moltiplichiamo ambo i membri per λ^{-1} ottenendo l'equazione:

$$\lambda^{-1}x = \lambda^{-1}A^{-1}\lambda x$$

la quale, sfruttando il fatto che il prodotto di una matrice per uno scalare è un'operazione commutativa, possiamo riscrivere come:

$$\lambda^{-1}x = \lambda^{-1}\lambda A^{-1}x$$

e diventa quindi:

$$\lambda^{-1}x = A^{-1}x$$

(2.3.1) Sappiamo quindi, che **se λ è un autovalore per A, allora λ^{-1} è un autovalore per A^{-1} .**

Ora, se *p* è la nostra approssimazione di λ (che vogliamo determinare), possiamo scrivere che:

$$Ax = px, x \neq 0$$

da cui:

$$(A - pI)x = 0 \quad \text{ovvero} \quad Ax - px = 0$$

Sostituendo la matrice col suo autovalore, abbiamo:

$$\lambda x - px = 0$$

che raccogliendo *x* può essere riscritto come:

$$(\lambda - p)x = 0$$

Dunque:

$$(A - pI)x = (\lambda - p)x$$

Sappiamo quindi che $(\lambda - p)$ è un autovalore della matrice $(A - pI)$ e dalla conclusione *(2.3.1)* che $(\lambda - p)^{-1}$ è un autovalore della matrice $(A - pI)^{-1}$.

Definiamo ora:

$$\mu = \frac{1}{\lambda - p}$$

Perciò quando l'approssimazione di *p* è buona (ovvero *p* è molto vicino a λ), la differenza $(\lambda - p)$ è molto piccola, pertanto il denominatore di μ diventa molto grande.
Posso quindi immaginare che se μ è molto grande, sia l'autovalore di massimo modulo per la matrice $(\lambda - p)^{-1}$.
Quindi possiamo applicare il metodo delle potenze visto nel paragrafo precedente, per ottenere un valore approssimato per μ. Pertanto, poiché:

$$\lambda = \frac{1}{\mu} + p$$

trovo un valore più accurato per λ.

Interpolazione polinomiale:

Motivazioni:

Molto spesso, in problemi matematici o addirittura nella costruzione stessa di metodi matematici di base, emerge la necessità di approssimare una funzione $f(x)$, definita attraverso una sua rappresentazione analitica oppure nota solo in alcuni punti $\{x_i\}$, con un'altra funzione, che chiameremo $f_n(x)$, di forma più semplice su cui si possa facilmente operare (derivare, integrare, ecc...).

Facciamo un'ipotesi: abbiamo eseguito delle misurazioni $\{yi\}$ (ad esempio di un fenomeno fisico, chimico, in campo sociale, zoologico, ecc...) ed abbiamo dei corrispondenti valori fissati $\{x_i\}$. Si vuole ora costruire un modello matematico (nel nostro caso una funzione) che descriva "sufficientemente* " bene il fenomeno in questione e che ci permetta di avere delle stime attendibili in quello che accade in punti x diversi dai nostri nodi $\{x_i\}$.

Per questo si utilizza l'interpolazione!

Che cos'è? Diamo una definizione matematica:

Dati $n+1$ punti distinti ed una funzione approssimante $g(x)$ diremo che g interpola i punti dati se:

$$g(x_i)=y_i \quad per \quad i=0,1,...,n$$

cioè se il grafico di g passa per i punti dati.

Se la funzione g è un polinomio parleremo di interpolazione polinomiale, se è una funzione razionale parleremo di interpolazione razionale, se g è trigonometrica parleremo di interpolazione trigonometrica etc... .

Facciamo degli esempi:

Se avessimo due punti: (x_0, y_0) e (x_1, y_1); e volessimo determinare la retta passante per entrambi, basterebbe risolvere il sistema:

$$\begin{cases} P_1(x_o) & = ax_0 + b = y_0 \\ P_1(x_1) & = ax_1 + b = y_1 \end{cases}$$

Analogamente, dati 3 punti (x_0, y_0), (x_1, y_1) e (x_2, y_2) possiamo determinare la

* vedremo poi che cosa intendiamo per sufficientemente.

parabola passante per tali punti risolvendo il sistema:

$$\begin{cases} P_2(x_o) & = ax_0^2 + bx_0 + c_0 = y_0 \\ P_2(x_1) & = ax_1^2 + bx_1 + c_1 = y_1 \\ P_2(x_2) & = ax_2^2 + bx_2 + c_2 = y_2 \end{cases}$$

Prima soluzione: Metodo dei coefficienti indeterminati

Più in generale, dati $n+1$ punti $(x_0, y_0), (x_1, y_1) \ldots (x_n, y_n)$, possiamo calcolare il polinomio passante per tali punti risolvendo il sistema:

$$\begin{cases} P_n(x_o) & = a_0 + a_1x_0 + a_2x_0^2 + a_3x_0^3 + \ldots + a_nx_0^n = y_0 \\ P_n(x_1) & = a_0 + a_1x_1 + a_2x_1^2 + a_3x_1^3 + \ldots + a_nx_1^n = y_1 \\ \vdots & \\ P_n(x_n) & = a_0 + a_1x_n + a_2x_n^2 + a_3x_n^3 + \ldots + a_nx_n^n = y_n \end{cases}$$

Si dimostra che se i nodi di interpolazione sono distinti tra loro (cioè: se $i \neq j$ allora $x_i \neq x_j$), il determinante della matrice dei coefficienti è diverso da 0 e il sistema ammette una ed una sola soluzione.

A tale proposito:
Teorema: Esiste uno ed un solo polinomio di grado n che assume valori y_i in corrispondenza degli $n+1$ punti distinti x_i con $i = 0, 1, \ldots, n$.

Quindi il polinomio interpolante di Lagrange è unico!

Abbiamo detto dunque che la matrice dei coefficienti del polinomio interpolante ha la seguente caratteristica: $det(A) \neq 0$ se e solo se per $i \neq j$ allora $x_i \neq x_j$.

Dunque la matrice dei coefficienti risulta essere una matrice di Vandermonde:

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ 1 & \ldots & \ldots & \ldots & \ldots \\ 1 & \ldots & \ldots & \ldots & \ldots \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{pmatrix}$$

Tale matrice ha però il difetto di essere fortemente mal condizionata (vedi mal condizionamento). Inoltre si nota facilmente anche che la risoluzione del sistema richiede un grande sforzo in termini di operazioni aritmetiche.

L'analisi di questo metodo ci è stato però utile per dimostrare l'esistenza e l'unicità del polinomio interpolante di grado n .

Ci rendiamo dunque conto che per trovare il polinomio interpolante è necessario cercare rappresentazioni alternative ben condizionate e, possibilmente, meno costose in numero di operazioni aritmetiche .

Seconda soluzione: Polinomi fondamentali di Lagrange

Una possibile soluzione al problema è quella di scrivere il polinomio interpolante mediante i polinomi di Lagrange (interpolazione di Lagrange):

Formula di interpolazione di Lagrange:

$$\mathrm{P}_n(x) = \sum_{k=0}^{n} l_{n,k}(x) y_k$$

$P_n(x)$ è il polinomio interpolante di Lagrange di grado n e $y_k = f-(x,k)$.

Esaminiamo ora il fattore $l_{n,k}(x)$:

$$l_{n,k}(x) = \prod_{\substack{i=0 \\ i \neq k}}^{n} \frac{x - x_i}{x_k - x_i} = \frac{(x - x_0)\dots(x - x_{k-1})(x - x_{k+1})\dots(x - x_n)}{(x_k - x_0)\dots(x_k - x_{k-1})(x_k - x_{k+1})\dots(x_k - x_n)}$$

Si nota facilmente che, per come sono definiti tali polinomi si ha che:

$$l_{n,k}(x_j) = \begin{cases} 1 & \text{se } j = k \quad \text{Perchè numeratore e denominatore sono uguali} \\ 0 & \text{se } j \neq k \quad \text{Perchè nel termine } (x_j - x_k) \text{ il denominatore si annulla} \end{cases}$$

In questo modo il polinomio di interpolazione $P_n(x)$, definito dalla matrice dei coefficienti (vedi sopra), viene espresso come combinazione lineare dei polinomi $\{l_{n,k}(x)\}$, ed i coefficienti di tale combinazione sono proprio le ordinate $\{y_i\}$.
I polinomi $\{l_{n,k}(x)\}$ sono detti polinomi fondamentali di Lagrange associati ai nodi $\{x_i\}$.

Vantaggi:

Il numero di operazioni aritmetiche diminuisce di molto.
Questo accade soprattutto nei casi in cui la funzione che l'interpolante deve "simulare" è per gran parte uguale a 0 o molto vicina a tale valore (ad esempio una curva gaussiana molto stretta).

Con funzioni del genere i nodi $\{x_i\}$ che ci vengono dati all'inizio sono quasi tutti uguali a 0 e gli unici elementi che partecipano alla sommatoria della formula (e che quindi danno origine ad una operazione aritmetica) sono quelli diversi da 0.

Svantaggi:

Questa rappresentazione è però soggetta a instabilità numeriche ed in particolare al fenomeno della cancellazione numerica.
Questo perché se si hanno 2 punti (chiamiamoli x_j ed x_{j+1}) molto vicini tra loro, nel calcolo di $l_{n,j}(x)$ e di $l_{n,j+1}(x)$ si avrà un denominatore molto vicino allo 0!

Terza soluzione: Polinomio interpolante di Lagrange espresso tramite rappresentazione di Newton (o metodo delle differenze divise)

Procediamo col riscrivere il polinomio interpolante nella seguente formula:

$$P_n(x)=y_0+(x-x_0)[x_0,x_1;f]+(x-x_0)(x-x_1)[x_0,x_1,x_2;f]+\ldots+(x-x_0)\ldots(x-x_n)[x_0,\ldots,x_n;f]$$

Per capire questa nuova rappresentazione del polinomio interpolante occorre prima capire cosa sono quegli elementi tra parentesi quadre; sono delle nuove quantità che chiameremo *Differenze divise*:

$$[a,b;f]=\frac{f(b)-f(a)}{b-a}$$

Questa (sopra) è la differenza divisa del primo ordine su due punti.

La differenza divisa del secondo ordine su tre punti è la seguente:

$$[a,b,c;f]=\frac{[b,c;f]-[a,b;f]}{c-a}$$

Quella di ordine 3 su 4 punti:

$$[a,b,c,d;f]=\frac{[b,c,d;f]-[a,b,c;f]}{d-a}$$

e così via.

Solo per maggiore chiarezza aggiungiamo la differenza divisa di ordine 0:

$$[a;f]=f(a)$$

e quella generale:

$$[a_0,\ldots,a_n;f]=\frac{[a_1,\ldots,a_n;f]-[a_0,\ldots,a_{n-1};f]}{a_n-a_0}$$

Notiamo dunque come costruire queste differenze divise sia un lavoro ricorsivo!

Facciamo vedere che, con questa rappresentazione, avendo 2 punti (x_0,y_0) e (x_1,y_1) si ottiene direttamente l'equazione della retta passante per questi 2 punti:

$$P_n=y=\frac{y_1+y_0}{x_1-x_0}(x-x_0)$$

Era comunque ovvio che utilizzando il polinomio interpolante con questi 2 punti si sarebbe ottenuta la retta passante per essi dato che è proprio quello che si cerca di ottenere con l'interpolazione; quello che salta all'occhio è però la facilità con cui ciò accade nonostante la formula di partenza sia una formula generale.

Vantaggi:

Il maggiore vantaggio di questo metodo consiste nel fatto che è possibile costruire ricorsivamente tutti i coefficienti a_0, a_1, … , a_n di $P_n(x)$. Ci possiamo rendere conto infatti che:

$$a_0=y_0 \qquad a_1=[x_0,x_1;f] \qquad a_2=[x_0,x_1,x_2;f] \qquad \cdots \qquad a_n=[x_0,\ldots,x_n;f]$$

Che sono le differenze divise presenti nella formula originale.

Ci rendiamo subito conto che per valori molto vicini (limite) dei 2 punti la differenza divisa del primo ordine equivale alla derivata prima di f :

$$[a,b;f]=\frac{f(b)-f(a)}{b-a}=f'(a)$$

Analogamente, in presenza di valori molto vicini, si ha che :

$$[a,b,c;f]=\frac{f''(a)}{2}$$

e procedendo:

$$[a,b,c,d;f]=\frac{f'''(a)}{3!}$$

Si può continuare all'infinito; in generale si ha:

$$[a,b,\ldots,n;f]=\frac{f^{(n)}(a)}{n!}$$

Ovviamente questo risolve il problema della cancellazione numerica che si era presentato con il precedente metodo (polinomi fondamentali di Lagrange); non è più un problema calcolare un polinomio interpolante quando si hanno dei nodi molto vicini (anzi, il calcolo è anche più diretto!).

Svantaggi:

Soprattutto dalle ultime argomentazioni illustrate si capisce che questa rappresentazione ha senso soltanto se la funzione da interpolare è derivabile.

Stima dell'errore (Resto dell'interpolazione):

Data una successione di punti ed una funzione interpolante vogliamo sapere quanto tale funzione si discosta dalla reale funzione di partenza in un certo punto x , cioè:

$$R_n(f,x)=\left|f(x)-P_n(x)\right|$$

L'espressione $R_n(f,x)$ viene chiamata resto o errore dell'interpolazione lagrangiana.

Possiamo di nuovo rappresentare tale quantità nella forma delle differenze divise:

$$R_n(f,x)=\left[x,x_0,\ldots,x_n;f\right](x-x_0)\ldots(x-x_n)$$

Ora, dall'analisi fatta prima (vedi sopra) sappiamo che per questa espressione esisterà un punto ξ tale che:

$$R_n(f,x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)\ldots(x-x_n)$$

Sappiamo dell'esistenza del punto ξ , ma non siamo in grado di stabilire dove si trovi. Benché questo fatto sia puramente teorico e quindi non può essere utilizzato dal punto di vista applicativo, ci permette di esplorare il seguente caso particolare:
Se la funzione $f\in\wp_n$ (l'insieme di tutti i polinomi di grado n), allora la sua derivata $n+1-esima$ sarà pari a 0 (zero) e di conseguenza:

$R_n(f,x)=0$ e quindi $P_n(x)=f(x)$

cioè l'interpolante andrà a coincidere con la funzione f (che ricordiamo essere un polinomio). Quindi se per interpolare un polinomio di grado n utilizziamo almeno $n+1$ punti la funzione risultante sarà il polinomio stesso.
Infatti sapevamo già da prima che per interpolare una retta (polinomio di grado 1) bastano 2 punti e per interpolare una parabola (polinomio di grado 2) bastano 3 punti ecc...; era lecito dunque aspettarci un risultato del genere!

Ora analizzeremo il comportamento dell'interpolante al crescere dei punti di interpolazione.

Convergenza dell'interpolazione:

Si dimostra che:

$$max\left|R_n(f,x)\right| \leq cE_n(f)(1+\lambda_n) \ , \ c \in \Re$$

dove

$$E_n(f) = \min_{P \in \wp_n} \max_{[x_0, x_n]} |f(x) - P(x)|$$

E' l'errore di migliore approssimazione uniforme. L'espressione di $E_n(f)$ in pratica esprime il fatto che si sceglie tra tutti i polinomi di grado n quello che si discosta meno da f e si prende l'errore massimo di interpolazione che si trova con quel polinomio.

Mentre:

$$\lambda_n = \max_{[x_0, x_n]} \sum_{k=0}^{n} |l_{n,k}(x)|$$

E' l' n-esima costante di Lebesgue.

Tanto più la funzione f è regolare, tanto più l'errore di interpolazione convergerà rapidamente a 0 (zero) per n che tende a ∞ (infinito).

In particolare:

$$se\ f \in C^0 \Rightarrow E_n(f)\ converge\ a\ 0;$$
$$se\ f \in C^r \Rightarrow E_n(f)\ converge\ a\ 0\ come\ 1/n^r;$$

C^0 indica che la funzione è continua;
C^n che la funzione è derivabile in ogni suo punto n volte.

Le quantità λ_n tendono ad infinito al crescere di n ed in particolare si ha che:

$$\lambda_n \geq c\dot{\log}(n)$$

Per bilanciare il cattivo comportamento delle quantità λ_n rispetto al buon comportamento dei $E_n(f)$ affinché il resto $R_n(f,x)$ tenda il più possibile velocemente a 0, bisogna operare delle scelte oculate sui punti di interpolazione x_i .

- $se\ \ x_i = zeri\ di\ Cebicev \Rightarrow \lambda_n \simeq \log(n)$ (Scelta ottimale)
- $se\ \ x_i = zeri\ di\ polinomi\ ortogonali \Rightarrow \lambda_n \simeq n^{\alpha}$
- $se\ \ x_i\ \ equidistanti \Rightarrow \lambda_n \simeq e^n$ (Scelta peggiore)

Quindi, facendo qualche esempio:

$se\ f \in C^1\ \ e\ \ x_i = zeri\ di\ Cebicev$ si avrà: $E_n(f) \to \frac{1}{n}$ e $(1+\lambda_n) \to \log(n)$

per cui:

$R_n(f,x) \to \frac{\log(n)}{n}$ e quindi (dato che n tende ad infinito) $R_n(f,x) \to 0$

Mentre,

$se\ f \in C^0\ \ e\ \ x_i\ \ equidistanti$ si avrà: $E_n(f) \to 0$ e $(1+\lambda_n) \to e^n$

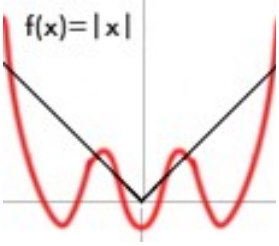
per cui con n tendente ad infinito:

$R_n(f,x) \to \infty$

Si vede dunque che la qualità dell'interpolazione dipende dal tipo di funzione da interpolare ed ancor più fortemente dalle scelte che si operano per i punti di interpolazione!

Un altro esempio:

Consideriamo la funzione $f = |x|$ e studiamo la convergenza dell'interpolazione scegliendo punti x_i equidistanti:



La funzione $f(x) \in C^0$ poiché come vediamo presenta un punto $(0\ ,\ 0)$ in cui la funzione non è derivabile; inoltre dato che i punti x_i scelti sono equidistanti, si avrà

che $\lambda_n \simeq e^n$ e quindi l'interpolante presenterà numerose oscillazioni (fenomeni di overflow ed underflow), mentre invece $E_n(f)$ tenderà a 0 (zero) molto lentamente. Per cui l'errore ($R_n(f,x)$) tenderà ad infinito piuttosto che ridursi all'aumentare dei punti di interpolazione.

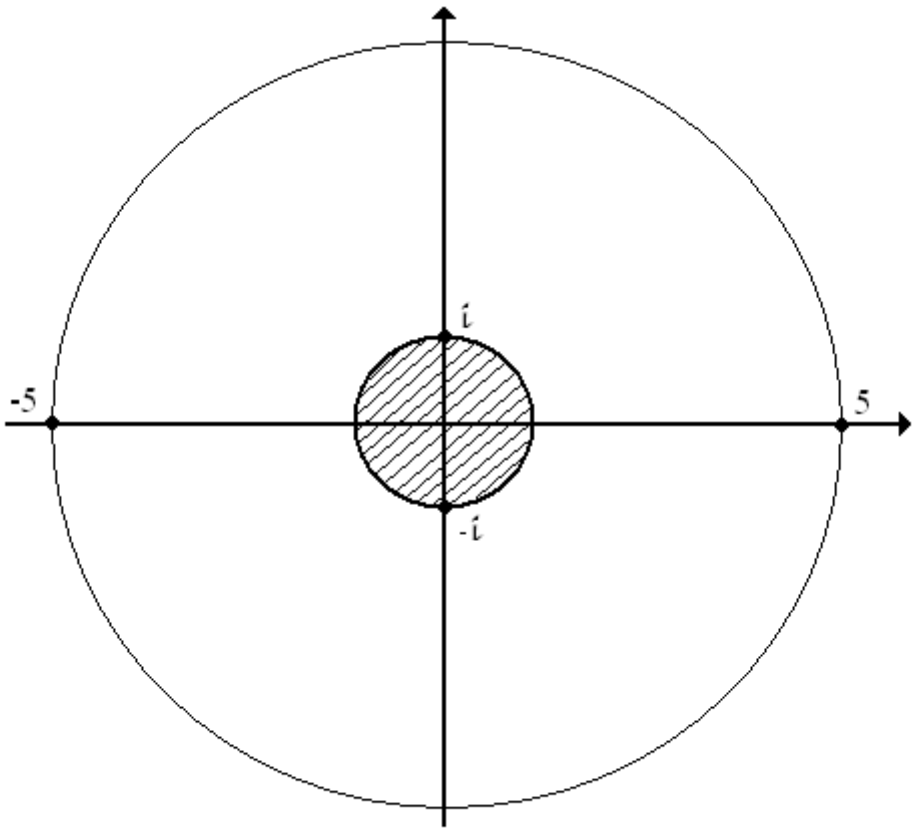
Fenomeno di Runge:

Esaminiamo la funzione

$$f(x)=\frac{1}{1+x^2} \quad \text{in } [-5\ ,\ 5]$$

Runge si accorse che per questa funzione (nonostante sia una funzione analitica, cioè derivabile infinite volte) la funzione interpolante non converge in tutto l'intervallo [-5 , 5] . Infatti se estendiamo la funzione al campo dei complessi si vede che il denominatore si annulla in $x=\pm 1$ dato che $i^2=-1$. Chiamiamo questi punti (quelli in cui $x=\pm 1$) punti di singolarità.

Quindi se i punti di interpolazione vengono scelti all'interno del piano complesso di raggio i l'interpolante converge sicuramente, mentre se si estende l'intervallo a cerchi più grandi che includono i punti di singolarità l'interpolante diverge.



Equazioni non lineari

Per equazioni non lineari si intendono equazioni della forma

$$f(x) = x^{\sqrt{x}} + \cos x + \log(x - 1)$$

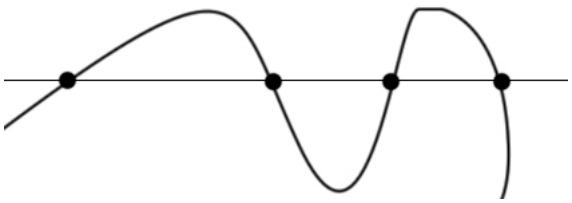
oppure funzioni apparentemente più semplici

$$f(x) = x^{1,98} + x + 1$$

che non sono risolvibili mediante i metodi dell'analisi classica.

É possibile però ottenere una approssimazione della soluzione di tali funzioni sotto alcune semplici ipotesi:

1. $f(x)$ è continua nell'intervallo $[a, b]$ in cui stiamo analizzando la funzione.
2. $f(a)f(b) < 0$: richiediamo che la funzione assuma segni alterni all'interno dell'intervallo; questo implica che esisterà almeno un punto della funzione che passa per l'asse x (zero della funzione).
3. $f'' \neq 0$: la funzione deve essere priva di flessi. In questo modo evitiamo situazioni in cui ci troviamo in presenza di radici multiple.



4. $[a, b]$ sufficientemente piccolo. Quest'ultimo requisito segue dallo sviluppo di Taylor, da cui sappiamo che l'errore dello sviluppo del polinomio è direttamente proporzionale alla dimensione dell'intervallo.

Sotto queste ipotesi è possibile applicare i cosiddetti metodi iterativi.

Metodi iterativi

Dato un punto iniziale x_0 detto ''punto di innesco'' si costruisce una successione di valori $x_1, x_2, \ldots, x_n$ mediante una funzione di interazione g tale che $x_{n+1} = g(x_n)$. La successione così costruita, al crescere di n, converge allo zero della funzione, o in altre parole

$$\lim_{n\to\infty} x_n = \overline{x} \quad \text{e} \qquad f(\overline{x}) = 0$$

Metodo di Newton-Raphson

In questo metodo la successione dei valori che approssima la soluzione della funzione, viene calcolata mediante la seguente formula:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

ovvero:

$$g(x) = x - \frac{f(x)}{f'(x)}$$

Poiché a denominatore compare la derivata prima di $f(x)$, per poter applicare questo metodo dovremo richiedere che, oltre alle ipotesi precedenti, valga anche che $f'(x) \neq 0$, cosi da evitare divisioni per 0.

Questo metodo ha convergenza quadratica, tuttavia risulta essere computazionalmente costoso poiché, ad ogni interazione, è necessario calcolare la derivata di $f(x)$.

Interpretazione geometrica

Il metodi di Metodo di Newton-Raphson è anche noto come ''metodo delle tangenti" perché per calcolare x_{n+1}, ad ogni iterazione si traccia la tangente alla curva nel punto di coordinate $(x_n, f(x_n))$ e il punto cercato è dato dall'intersezione della tangente con l'asse delle x.

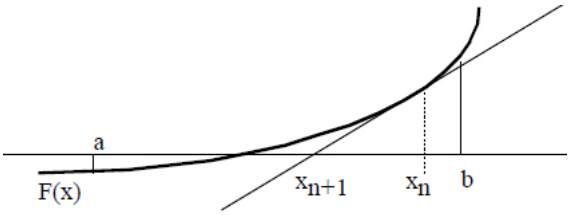


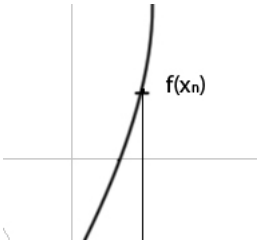
Figura 1: Metodo di Newton-Raphson

Test di convergenza (criterio di stop)

Cerchiamo ora di capire a quale iterazione del metodo di Newton-Raphson fermarci per ottenere una buona approssimazione dello zero della funzione che stiamo esaminando.

Ad una prima analisi potremo decidere di fermarci nel momento in cui l' n-esimo valore calcolato sia più piccolo di un certo valore fissato (es. $\frac{1}{1000}$), ovvero

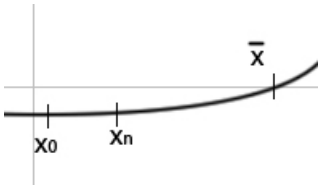
$$|f(x_n)| \leq \varepsilon$$



Tuttavia tale criterio non risulta essere efficace nei casi in cui, ad esempio, la funzione cresca rapidamente . Per quanto visto quindi, si potrebbe pensare di fermare il metodo non la differenza tra due punti calcolati sia piccola cioè:

$$|x_{n+1} - x_n| < \varepsilon$$

Purtroppo anche questo criterio non risulta essere ottimale in quanto nel caso di funzioni che crescono lentamente rischieremo di fermarci troppo presto come mostrato dalla figura.



Proprio per escludere questi casi limite, si dimostra che un ottimo criterio di stop è dato dalle seguenti due formule (che devono essere entrambe verificate)

$$|x_{n+1} - x_n| < \varepsilon$$
$$|x_{n+1} - x_n| < \varepsilon|x_n|$$

Quadratura numerica

Le formule di quadratura numerica permettono di ottenere un'approssimazione del valore degli integrali di funzioni $f(x)$ quando tali funzioni risultano difficili da calcolare mediante i metodi classici dell'analisi matematica. In generale

$$\int_a^b f(x)\ dx \cong \sum_{i=1}^{n} w_i f(x_i)$$

dove gli x_i sono nodi e i w_i sono pesi. Si dimostra che se al crescere di n $\sum |w_i| \leq k$ la quadratura converge, altrimenti, se la somma dei pesi non è limita, diverge.

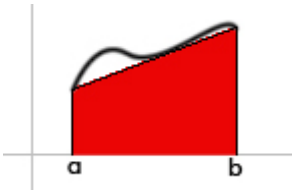
Formule di Newton-Cotes

Si applicano scegliendo nodi equidistanti. Dato che al crescere di n il modulo della somma dei pesi w_i tende ad infinito, queste formule vengono generalmente applicate su due o tre punti massimo.

Formula dei trapezi

$$\int_a^b f(x)\ dx \cong \frac{b-a}{2}\ [f(a)+f(b)] \quad \text{dove } h = b-a \tag{1}$$

La formula approssima l'integrale all'area del trapezio al di sotto del segmento $b-a$ e quindi in molti casi fornisce una stima poco accurata.



Formula di Cavalieri-Simpson

$$\int_a^b f(x)\ dx \cong \frac{h}{3}\ [f(a)+4f(M)+f(b)] \tag{2}$$

dove M rappresenta il punto medio tra a e b.

Formule composite

Proprio a causa del comportamento instabile al crescere dei punti utilizzati, abbiamo detto che è preferibile utilizzare le formule di Newton-Cotes per due o tre punti al massimo. Tuttavia, dovendo utilizzare obbligatoriamente nodi equidistanti, si può pensare di suddividere l'intervallo di applicazione $[a,b]$ in intervalli più piccoli di uguale ampiezza e su ciascuno di questi applicare le formule precedenti, cioè:

$$\int_a^b f(x)\ dx = \sum_{i=0}^{n-1} \int_{a_i}^{a_i+1} f(x)\ dx$$

A questo punto la formula in (1) applicata su più intervalli diventa:

$$\int_a^b f(x)\ dx \cong \frac{h}{2}\ [f(a) + f(b) + \sum_{i=0}^{n-1} f(a+ih)]$$

mentre la formula (2) diventa:

$$\int_a^b f(x)\ dx \cong \frac{h}{3}\ [f(a) + 2\sum_{i=1}^{n-1} f(a+2ih) + h\sum_{i=1}^{n-1} f(a+(2i-1)h) + f(b)]$$

Anche in questo caso le formule presentate risultano essere molto semplici ma poco accurate.

Formule di Gauss

A differenza delle formule precedenti, le formule di Gauss hanno il vantaggio di essere convergenti (questo vuol dire che al crescere del numero dei nodi utilizzati $\sum |w_i| \leq k$, o in altre parole la somma dei pesi è limitata); di contro però la loro costruzione è leggermente più complicata.

Si introduce una funzione peso $w(x)$ tale che

1. $w(x) \geq 0$
2. $\int_a^b w(x)\ dx = 1$

A partire da $w(x)$ si costruisce una successione di polinomi P_0 di grado 0, P_1 di grado 1, ..., P_n di grado n (successione di polinomi ortogonali) tale che:

$$\forall n, m \int_a^b w(x)P_n(x)P_m(x)\ dx = \begin{cases} 0, & \text{se } n \neq m \\ \neq 0, & \text{se } n = m \end{cases}$$

Inoltre, se l'integrale risulta essere uguale ad 1, diremo che la successione è ortonormale.

Funzioni peso

Di seguito vengono elencate alcune delle funzioni peso più note in letteratura.

funzione $w(x)$	intervallo di validità	nome
1	$(-1,1)$	polinomi di Legendre
$\frac{1}{\sqrt{1-x^2}}$	$(-1,1)$	polinomi di Cebicev di prima specie
$\sqrt{1-x^2}$	$(-1,1)$	polinomi di Cebicev di seconda specie
e^{-x}	$(0,+\infty)$	polinomi di Laguerre
e^{-x^2}	$(-\infty,+\infty)$	polinomi di Hermite

Costruzione

Sia P_n l'n-esimo polinomio ortogonale in $[a,b]$. Si individuano gli zeri (o radici) di tale polinomio rispetto alla funzione peso w, cioè quei punti x_i tali che $P_n(x_i)=0$ (P_n ha radici reali e distinte in $[a,b]$). Con questi punti si dimostra che

$$\int_a^b w(x)f(x)\,dx \cong \sum_{i=1}^{n} w_i f(x_i)$$

Le formule di Gauss hanno pesi positivi, cioè si dimostra che $w_i > 0$ per $i=0,\ldots,n$; questo comporta che $\sum |w_i| = \sum w_i = 1$ in quanto tali formule sono esatte per $f=1$ e sono quindi convergenti per quanto detto fino ad ora.

Per individuare gli x_i e i w_i si usa solitamente un software (QUADPACK) che a partire da $[a,b]$ e w fornisce in maniera automatica i valori cercati.

Esempio

Calcoliamo l'integrale seguente mediante le formule di quadratura di Gauss:

$$\int_{-1}^{1} \frac{x^5}{\sqrt{1-x^2}}\,dx$$

Scegliendo $w(x) = \frac{1}{\sqrt{1-x^2}}$ (polinomio di Cebicev di prima specie) si vede immediatamente che

$$\int_{-1}^{1} \frac{x^5}{\sqrt{1-x^2}}\,dx \cong \sum_{i=1}^{n} w_i x_i^5$$

In questo caso inoltre si dimostra che $w_i = \dfrac{\pi}{n}$ e che $x_i = \cos\dfrac{\pi}{2n}(2i+1)$.

In generale, volendo applicare una formula di quadratura per calcolare l'integrale di una funzione $g(x)$, si può moltiplicare e dividere l'integrando

per una funzione peso opportuna e quindi applicare la formula di quadratura come mostrato di seguito:

$$\int_a^b g(x)\,dx = \int_a^b \frac{g(x)\sqrt{1-x^2}}{\sqrt{1-x^2}}\,dx = \int_a^b \frac{f(x)}{\sqrt{1-x^2}}\,dx$$

dove $f(x) = g(x)\sqrt{1-x^2}$.

Nota

Se la formula integranda $f \in P_{2n-1}$ (ovvero è un polinomio di grado al più $2n-1$ allora la formula di quadratura è esattamente uguale al valore dell'integrale cioè

$$\int_a^b w(x)f(x)\,dx = \sum_{i=1}^{n} w_i f(x_i)$$

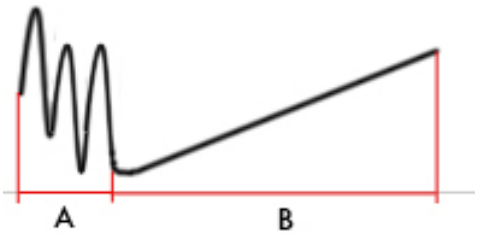
In questo caso si dice che la formula di quadratura ha grado di precisione (o di esattezza) $2n-1$. Quindi nell'esempio precedente, poiché $x^5 \in P_{2n-1}$, possiamo sostituire il simbolo $\cong$ con il simbolo di $=$ e quindi

$$\int_{-1}^{1} \frac{x^5}{\sqrt{1-x^2}}\,dx = \sum_{i=1}^{n} w_i x_i^5 = \frac{\pi}{n}\sum_{i=1}^{n} \cos^5 \frac{\pi}{2n}(2i+1)$$

Si dimostra che $2n-1$ è il massimo grado di precisione possibile per formule di quadratura di questo tipo e quindi le formule gaussiane sono ottimali.

Strategie di quadratura automatica

I software di quadratura automatica come il QUADPACK, cercano di applicano formule viste fino ad ora in maniera combinata al fine di ottenere dei valori che si avvicinano il più possibile ai valori esatti degli integrali.



Supponiamo di voler calcolare l'integrale della funzione rappresentata in figura. In questo caso il software tenderà a suddividere il grafico in due parti: la parte in B più regolare verrà calcolata utilizzando pochi punti mediante le formule di Gauss; la parte in A invece, che presenta un andamento meno regolare, richiederà l'utilizzo di più punti e l'uso delle formule di Newton-Cotes .

Approssimazione

1. Introduzione

Supponiamo che $f(x)$ sia una serie o una funzione molto complicata, in molti casi, può essere utile sostituirla con una più semplice che sia più facile da trattare (es. polinomio, funzione trigonometrica, etc). Data quindi una funzione $f(x)$, si utilizza una funzione $g(x)$ che approssima la funzione di partenza. Vogliamo sapere quanto $g(x)$ si scosta da $f(x)$, cioè quanto è l'errore dovuto all'approssimazione:

$f-g=?$

2. Norma di Chebyshev

Per valutare la distanza dell'approssimazione dalla funzione originale si introduce la Norma di Chebyshev (o Norma infinito o del massimo) definita come:

Definizione:
Sia $f(x)$ una funzione continua in un intervallo chiuso e limitato [a,b] , e sia $g(x)$ una funzione continua in [a,b], si dice Approssimazione di Chebyshev della funzione $g(x)$ rispetto alla funzione $f(x)$ la norma:

Se $f \in C$ $\qquad \|f-g\|_\infty = max_{[a,b]}|f(x)-g(x)|, \qquad$ con $f,g \in C$

Un esempio di successione di polinomi approssimanti una funzione continua è dato dai *polinomi di Berstein*. I polinomi di Berstein $B_n(x)$ relativi ad una funzione $f \in C^0$ in [0,1] convergono uniformemente in [0,1] alla $f(x)$.

3. Teorema di Weierstrass

Il problema che si presenta nello studio dell' *approssimazione polinomiale* è il seguente:

Data una funzione f è sempre possibile costruire una successione di polinomi $\{P_n(x)\}$ che converga uniformemente in [a,b], cioè tale che $\|P_n-f\|_\infty \to 0$?

La soluzione a questo problema è data dal *Teorema di Weierstrass*:

Assegnata una funzione continua in un intervallo chiuso e limitato [a,b] esiste almeno una successione di polinomi $P_n(x)$ convergente uniformemente verso $f(x)$ in [a,b].Una qualsiasi successione di polinomi che converge uniformemente a si chiama *successione di polinomi approssimanti* .

Va inoltre osservato che esistono più successioni di polinomi con tale proprietà (es. i polinomi di Bernstein, usati anche in grafica per le loro proprietà mimiche).

Se $f \notin C$, ma $\int_a^b f^2(x)dx < \infty$,

si può comunque trovare una buona approssimazione: vogliamo sostituire la f, che è difficilissima, con una funzione g più semplice, di cui sappiamo fare le derivate, e poi valutare l'errore che viene commesso.

4. Norma quadratica

Vogliamo trovare quindi la cosiddetta Norma quadratica che è definita come:

Definizione:
Sia $f(x)$ una funzione definita, e generalmente continua, in un intervallo chiuso e limitato [a,b] e di quadrato integrabile e sia $g(x)$ una successione di funzioni anch'esse generalmente continue e di quadrato integrabile in [a,b], si chiama Approssimazione in media quadratica di rispetto ad la norma:

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

Nelle funzioni di *quadrato integrabile* non possiamo calcolare il massimo (come quella definita nella norma di Chebyshev) poiché f non è continua.
Un risultato importante ci dice che esistono più funzioni g_k continue in [a,b] che formano un sistema ortonormale e a partire da queste ne costruiamo la formula:

$$f_n(x) = \sum_{k=1}^{n} c_k g_k(x)$$

si dimostra che questa f_n ha una proprietà importante: la quantità

$$\mu_n = \sqrt{\int_a^b [f(x) - \sum_{k=1}^{n} c_k g_k(x)]^2 dx}$$

è minima se i coefficienti c_k sono uguali ai coefficienti a_k (coefficienti di Fourier):

$$c_k = a_k \text{ , dove } a_k = \int_a^b f(x) g_k(x) dx$$

5. I polinomi di Chebyshev

Diamo ora la definizione dei polinomi ortogonali di Chebyshev,molto noti nel calcolo numerico e in vari contesti della matematica.

La costruzione dei polinomi di Chebyshev viene eseguita con una semplice formula: la *Relazione di ricorrenza a tre termini*.

Secondo quest'ultima basta conoscere due termini per ricavare il terzo:

posto $T_{0(x)}=1$, $T_{1(x)}=x$,
risulta $T_{n+1}(x)=2xT_{n(x)}-T_{n-1}(x)$,

ottenendo così polinomi T_{n+1} di grado $n+1$.

5.1 Zeri di Chebyshev

Gli zeri di Chebyshev sono dati da semplici espressioni,sono quantità reali e distinte l'una dall'altra(e questa è un'informazione importante per varie applicazioni):

$x_k=cos\frac{\pi}{2n}(2k+1)$, con *k=0,.....,n*

Si vuole ora dimostrare che i polinomi $T_n(x)$ costituiscono, nell'intervallo (-1,1) un sistema di funzioni ortogonali rispetto alla funzione peso $\frac{1}{\sqrt{1-x^2}}$:

Si può dimostrare che:

$$\int_{-1}^{+1}\frac{T_{m(x)}T_{n(x)}}{\sqrt{1-x^2}}dx \quad \begin{cases} = 0 & \text{se } n\neq m \\ = \frac{\pi}{2} & \text{se } n>0 \\ = \pi & \text{se } n=0 \end{cases}$$

che è proprio la condizione di **ortogonalità**.

Notiamo poi che il sistema di polinomi U_n , *n= 0,1,...,* con

$$U_0(x)=\frac{T_0(x)}{\sqrt{\pi}}$$

$$U_n(x)=T_n(x)\frac{\sqrt{2}}{\sqrt{\pi}}$$

è ortonormale in (-1,1) rispetto alla funzione peso $\frac{1}{\sqrt{1-x^2}}$; infatti è possibile dimostrare che:

$$\int_{-1}^{+1} \frac{Um(x)U_n(x)}{\sqrt{1-x^2}}dx \quad \begin{cases} =0 \text{ se } n \neq m \\ =1 \text{ se } n = m \end{cases}$$

Questi polinomi trovano applicazione nella *quadratura Gaussiana* (vedi paragrafo) e nella *quasi migliore apposimazione*.(vedi paragrafo 6.1)

6. Il problema della migliore approssimazione

Se si assegna un polinomio che approssima $f(x)$ in un intervallo chiuso e limitato [a,b], a questo corrisponde una ben determinata approssimazione data da:

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

ma se si fissa solo il grado del polinomio l'approssimazione varia al variare dei coefficienti.
Ora ci chiediamo se assegnata una funzione continua $f(x)$, esiste un polinomio $P_n(x)$ di grado non maggiore di n , per il quale l'approssimazione assuma il valore minimo.
Un polinomio $P_n(x)$ che realizzi la migliore approssimazione si chiama *polinomio di migliore approssimazione*.

Teorema: Assegnata una funzione continua $f(x)$ nell'intervallo chiuso e limitato [a,b] ad ogni intero positivo n corrisponde uno ed un solo polinomio $P_n(x)$ di grado non maggiore di n che realizza la migliore approssimazione di $f(x)$ in [a,b].

Questo teorema assicura l'esistenza e l'unicità del polinomio $P_{n(x)}$. Si può dire, quindi, che la successione $P_n(x)$ è quella che nell'intervallo (a,b) converge più rapidamente verso $f(x)$ .Purtroppo, però, la determinazione del polinomio di migliore approssimazione, a parte che per pochi casi particolari, è estremamente laboriosa.

6.1. Polinomio di quasi migliore approssimazione

Parliamo quindi di polinomi che hanno la proprietà di essere dei *buoni approssimanti*. Un modo efficace per costruirli è quello di assumere come polinomio approssimante il polinomio che minimizza l'approssimazione in media quadratica rispetto ad una opportuna funzione peso.

Considero:

$$f_n(x) = \frac{C_0 T_0}{\sqrt{\pi}} + \sum_{k=1}^{n} c_k T_{k(x)} \frac{\sqrt{2}}{\sqrt{\pi}}$$

Successione con polinomi di Chebyshev

Cioè scelgo la f_n come somma dei polinomi di Chebyshev e realizzo la migliore approssimazione in media quadratica dove:

$$c_0 = a_0 = \frac{1}{\sqrt{\pi}} \int_{-1}^{+1} f_{(x)} T_{0(x)}\, dx$$

$$c_k = a_k = \frac{\sqrt{2}}{\sqrt{\pi}} \int_{-1}^{+1} f_{(x)} T_{k(x)}\, dx$$

, per *k=1,.....,n*

Si dimostra che tale f_n è una buona approssimazione di f,se f è una funzione continua in [-1,1]. Quindi f_n non realizza la migliore approssimazione uniforme (è stata costruita per ottenere la migliore approssimazione in media quadratica), ma da buoni risultati, pertanto viene chiamata polinomio di quasi migliore approssimazione uniforme.

6.2 Approssimazione ai minimi quadrati

Supponiamo di avere $n+1$ coppie di punti distinti (x_0,y_0),....., (x_n,y_n) e di voler approssimare la successione di questi punti con la retta $y = a_0 + a_1 x$ tale che lo scostamento tra i punti e la retta stessa sia minimo. Questo equivale a calcolare i valori di a_0 e a_1 che minimizzano lo scostamento, ovvero

$$min_{a_0,a_1} S = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_1)^2$$

dove:

$$S = \sum_{i=0}^{n} (y_i^2 + a_0^2 + a_1^2 x_i^2 - 2a_0 y_i - 2a_1 x_i y_i + 2a_0 x_i y_i)$$

Dall'analisi sappiamo che il minimo di una funzione a due variabili si ottiene derivando rispetto alle singole variabili, ponendo cioè:

$$\frac{\delta S}{\delta a_0} = -2\sum_{i=1}^{n}(y_i - a_0 - a_1 x_i) = 0 \qquad \text{e} \qquad \frac{\delta S}{\delta a_1} = -2\sum_{i=1}^{n}(y_i - a_0 - a_1 x_i)x_i = 0$$

poichè le due quantità sono nulle, posso semplificare il 2.

$$\begin{cases} N a_0 + \sum x_i a_1 = \sum y_i \\ \sum x_i a_0 + \sum x_i^2 a_1 = \sum x_i y_i \end{cases}$$

La soluzione di tale si sistema sarà quindi:

$$a_0 = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{N\sum x_i^2 - (\sum x_i)^2} \qquad a_1 = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{N\sum x_i^2 - (\sum x_i)^2}$$

in questo modo ricaviamo i coefficienti a_0 e a_1 del polinomio di primo grado e possiamo tracciare la retta che passa per i punti, detta retta ai minimi quadrati.

Analogamente possiamo approssimare la successione dei punti con un polinomio di grado *M* della forma:

$$y = a_0 + a_1 x + a_2 x^2 + \ldots\ldots + a_M x^M$$

per fare questo basta calcolare i coefficienti $a_0,\ldots\ldots,a_M$ di tale equazione:

$$S = \sum_{i=1}^{n}(y_i - a_0 - a_1 x_1 - a_2 x_i^2 - \ldots\ldots\ldots - a_M x_i^M)^2$$

ovvero:

$$\begin{cases} \frac{\delta S}{\delta a_0} = -\sum_{i=1}^{n} 2(y_i - a_0 - a_1 x_i - \ldots\ldots\ldots - a_M x_i^M) = 0 \\ \frac{\delta S}{\delta a_1} = -\sum_{i=1}^{n} 2(y_i - a_0 - a_1 x_i - \ldots\ldots\ldots - a_M x_i^M)x_i = 0 \\ \frac{\delta S}{\delta a_2} = -\sum_{i=1}^{n} 2(y_i - a_0 - a_1 x_i - \ldots\ldots\ldots - a_M x_i^M)x_i^2 = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{\delta S}{\delta a_M} = -\sum_{i=1}^{n} 2(y_i - a_0 - a_1 x_i - \ldots\ldots\ldots - a_M x_i^M)x_i^M = 0 \end{cases}$$

Se imponiamo che queste derivate siano uguali a 0, alla fine otteniamo un sistema di equazioni in cui *H* è il vettore contenente la soluzione:

H a = b con $a=(a_0,a_1,.....,a_M)^M$ $b=(\sum y_i,\sum x_i y_i,\sum x_i^2,y_i,......,\sum x_i^M y_i)$

dove *b* è stato ricavato dalle equazioni normali del seguente sistema:

$$\begin{cases} Na_0+\sum x_i a_1+\sum x_i^2 a_2+.........+\sum x_i^M a_M=\sum x_i \\ \sum x_i a_0+\sum x_i^2 a_1+\sum x_i^3 a_2+.........+\sum x_i^{M+1} a_M=\sum x_i y_i \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum x_i^M a_0+\sum x_i^{M+1} a_1+\sum x_i^{M+2} a_2+.........+\sum x_i^{2M} a_M=\sum x_i^M y_i \end{cases}$$

$$H=\begin{pmatrix} N & \sum x_i,\ldots\ldots\ldots, & \sum x_i^M \\ \sum x_i & \sum x_i^2,\ldots\ldots\ldots, & \sum x_i^{M+1} \\ \sum x_i^2 & \sum x_i^3,\ldots\ldots\ldots, & \sum x_i^{M+2} \\ \ldots\ldots\ldots & & \\ \sum x_i^M & \sum x_i^{M+1},\ldots\ldots\ldots, & \sum x_i^{2M} \end{pmatrix}$$

In generale è possibile estendere questo metodo ad un qualsiasi polinomio di grado *r* , tuttavia tale approccio è fortemente sconsigliato per polinomi di grado superiore a 2,in quanto porta alla risoluzione di un sistema di equazioni lineari malcondizionato e conseguentemente ad una soluzione poco affidabile.

Malcondizionamento

Partiamo con un esempio:
Consideriamo i 2 sistemi di equazioni

$$\begin{cases} x - y = 1 \\ x - 0,99999y = 0 \end{cases} \qquad \begin{cases} x - y = 1 \\ x - 1,00001y = 0 \end{cases}$$

I 2 sistemi differiscono di poco (solo un coefficiente nelle seconde equazioni), tuttavia le soluzioni dei 2 sistemi differiscono di molto.
Infatti la prima soluzione è: x = -99.999 y = -100.000
La seconda soluzione è: x = 100.001 y = 100.000

Questa grande differenza è dovuta al fatto che le 2 equazioni in ognuno dei sistemi rappresentano rette quasi parallele; due rette parallele si incontrano all'infinito, essendo queste quasi parallele si incontreranno in un punto molto lontano dall'origine degli assi; la differenza nei 2 sistemi è nel coefficiente della seconda equazione, nel primo è minore di uno, nel secondo è maggiore, quindi nel primo caso le 2 rette si incontreranno molto lontano nel terzo quadrante e nel secondo caso si incontreranno molto lontano nel primo quadrante.

Un tipo di problema come questo, dove ad una piccola variazione dei dati di input, corrisponde una grande variazione dei valori della soluzione, viene definito problema malcondizionato.

Norma

Per definire l'indice di mal condizionamento di una matrice bisogna introdurre la nozione di norma. La norma, definita con $\|\ \|$, è una funzione reale non negativa che soddisfa le seguenti proprietà:

- $||A|| > 0 \qquad , A \neq 0$
- $||A|| = 0 \quad \Leftrightarrow \quad A = 0$
- $||\alpha \cdot A|| \ = \ |\alpha| \cdot ||A|| \ , \quad \alpha \in \Re$
- $||A + B|| = ||A|| + ||B||$ (diseguaglianza triangolare)

Ci sono vari ti pi di norma:

- $||A||_\infty = \max_i \sum_j |a_{ij}|$ (norma infinito)
- $||A||_1 = \max_j \sum_i |a_{ij}|$ (norma 1)
- $||A||_2 = \max_i |\lambda_i|$ (norma quadratica o norma 2)
- $||A||_F = \sqrt{\sum_{i,j} a_{ij}^2}$ (norma di Frobenius)

Tutte si equivalgono a meno di una costante, quindi è indifferente quale usare.

Come vedere se un sistema è malcondizionato

Consideriamo il sistema di equazioni

(1) $Ax = b$

b è caratterizzato (sicuramente) da un errore di misurazione
quindi lo prendiamo come $b + \Delta b$.
In corrispondenza quindi il sistema diverrà

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

Essendo $Ax = b$ rimane $A\Delta x = \Delta b$, quindi

(3) $\Delta x = A^{-1}\Delta b$

Partendo dall'ultima uguaglianza (3) e passando alla norma, otteniamo:

(4) $||\Delta x|| = ||A^{-1}\Delta b|| \leq ||A^{-1}|| * ||\Delta b||$

e dalla (1) otteniamo:

(5) $||b|| = ||Ax|| \leq ||A|| * ||x||$

Mantenendo l'uguaglianza (4) e dalla (1) possiamo dividere $||\Delta x||$ per $||Ax||$ e $||A^{-1}\Delta b||$ per $||b||$
In modo da avere

(6) $$\frac{||\Delta x||}{||Ax||} = \frac{||A^{-1}\Delta b||}{||b||}$$

Sfruttando poi le proprietà della norma avremo

(7) $$\frac{||\Delta x||}{||A|| * ||x||} \leq \frac{||A^{-1}|| * ||\Delta b||}{||b||}$$

Moltiplicando entrambi i fattori per la norma di A giungeremo a

(8) $$\frac{||\Delta x||}{||x||} \leq ||A|| \cdot ||A^{-1}|| \cdot \frac{||\Delta b||}{||b||}$$

Dove il termine $||A|| * ||A^{-1}||$ è detto **indice di condizionamento** della matrice.

L'errore è sempre amplificato

A dimostrazione di ciò si prenda la seguente uguaglianza

(9) $$||I|| = 1 = ||A \cdot A^{-1}|| \leq ||A|| \cdot ||A^{-1}||$$

che fa vedere come il coefficiente di condizionamento che si affianca all'errore sia sempre maggiore di 1.

Nota che

(10) $\frac{||\Delta b||}{||b||}$ rappresenta l'errore relativo sul termine noto e

(11) $\frac{||\Delta x||}{||x||}$ rappresenta l'errore relativo sulla soluzione del sistema.

Pertanto dalla (8) e dalla (9) segue che l'errore relativo sulla soluzione viene sempre amplificato.

Esempio di problema condizionato

Scelta di una retta che approssima un insieme di punti

Abbiamo una serie di punti (X_i , Y_i) con $i = 1.....N$
E una retta $Y = a_0 + a_1X$

Se voglio che la retta si scosti poco da tutti i punti dovrò soddisfare $\sum Y_i - a_0 - a_1X_i$, cioè trovare a_0 e a_1 tali che sia minimizzata la sommatoria $S = \sum (Y_i - a_0 - a_1X_i)^2$ (tecnica dei minimi quadrati).
Per trovare i minimi si passa alle derivate, quindi

$\frac{dS}{da0} = -2\sum (Y_i - a_0 - a_1X_i) = 0$ è la derivata rispetto ad a_0

$\frac{dS}{da1} = -2X_i(Y_i - a_0 - a_1X_i) = 0$ è la derivata rispetto ad a_1

Sviluppando le sommatorie otteniamo

$\frac{dS}{da0} = -Na_0 - \sum X_ia_1 + \sum Y_i$ quindi $\sum Y_i = Na_0 + \sum X_ia_1$

$\frac{dS}{da1} = -\sum X_ia_0 - \sum X^2{}_ia_1 + \sum X_iY_i$ quindi $\sum X_iY_i = \sum X_ia_0 + \sum X^2{}_ia_1$

Trasformando il nostro sistema in matrice otteniamo

$$\begin{matrix}\sum Y_i \\ \sum X_iY_i\end{matrix} = \begin{pmatrix} N & \sum X_i \\ \sum X_i & \sum X^2{}_i \end{pmatrix} * \begin{matrix} a_0 \\ a_1 \end{matrix}$$

Quindi

$$a_0 = \frac{\begin{vmatrix} \sum Y_i & \sum X_i \\ \sum X_i Y_i & \sum X^2{}_i \end{vmatrix}}{\begin{vmatrix} N & \sum X_i \\ \sum X_i & \sum X^2{}_i \end{vmatrix}} = \frac{\sum X^2{}_i * \sum X_i - \sum X_i * \sum X_i Y_i}{N * \sum X^2{}_i - \sum X_i * \sum X_i}$$

$$a_1 = \frac{\begin{vmatrix} N & \sum Y_i \\ \sum X_i & \sum X_i Y_i \end{vmatrix}}{\begin{vmatrix} N & \sum X_i \\ \sum X_i & \sum X^2{}_i \end{vmatrix}} = \frac{N * \sum X_i Y_i - \sum Y_i * \sum X_i}{N * \sum X^2{}_i - \sum X_i * \sum X_i}$$

Questi 2 valori trovati saranno utili al loro scopo soltanto se la matrice $H = \begin{pmatrix} N & \sum X_i \\ \sum X_i & \sum X^2{}_i \end{pmatrix}$ ha un basso indice di condizionamento.

Nelle matrici trovate come risultato dell'uso dei minimi quadrati l'indice di condizionamento dipende dalla dimensione delle matrici stesse; nel nostro caso dipende quindi dal numero di coefficienti che dobbiamo trovare.

Equazioni differenziali (del primo ordine)

1.Introduzione

In matematica e in molti problemi scientifici riveste un ruolo cruciale la risoluzione di equazioni differenziali del primo ordine (ad es. esse modellano vari fenomeni evolutivi).Si parla di equazioni differenziali del primo ordine perchè la funzione incognita y appare derivata proprio al primo ordine.

Dato il sistema:

$$\begin{cases} y'(x) = f(x, y(x)) \\ y(x_0) = y_0 \end{cases} \quad (1^*)$$

vogliamo calcolare il valore della funzione y nell'intervallo (x_0, x_n). Tranne che per casi molto semplici, tale sistema non può essere risolto mediante i metodi classici dell'analisi. Tuttavia mediante i metodi che verranno illustrati di seguito è possibile calcolare una soluzione approssimata di tale sistema.

In particolare, suddividiamo l'intervallo (x_0, x_n) in intervallini uguali di ampiezza $h = x_i + 1 - x_i$ (tale ampiezza è detta anche passo di integrazione dell'equazione differenziale) e cerchiamo quindi la soluzione approssimata per ciascuno dei punti x_i , ovvero calcoliamo per ogni punto il valore di $y(x_i)$.

1.2 Metodo di Eulero

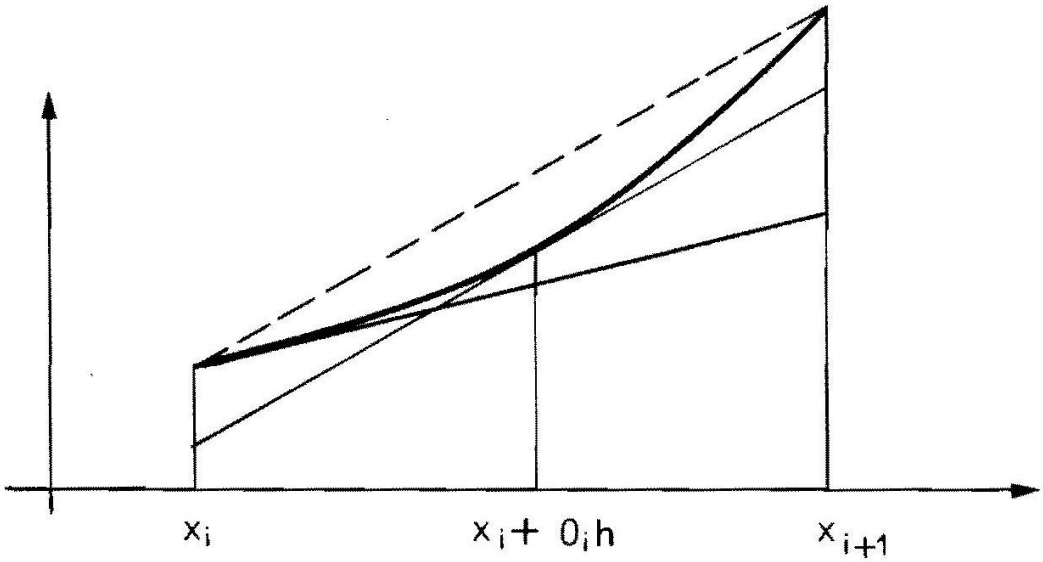
Dall'analisi sappiamo che vale la seguente relazione:

(1) $y(x_{i+1}) - y(x_i) = y'(\xi) + (x_{i+1} - x_i) = h y'(\xi)$ con $x_i < \xi < x_{i+1}$

Sappiamo dell'esistenza di questo punto, ma non sappiamo esattamente dove si trova. Il metodo di Eulero approssima la derivata di f in ξ con la derivata di f in x_i, quindi:

(2) $y'(\xi) \simeq y'(x_1)$

La seguente figura mostra che questa approssimazione è grossolana in quanto non è vero che $y'(\xi)$ è simile a $y'(x_i)$.



Da questo momento in poi a causa dell'approssimazione (2) stiamo introducendo un metodo numerico (cioè basato su un'approssimazione) per risolvere il problema (1*).

Da questo momento in poi la lettera y denoterà la soluzione approssimata e si userà la notazione abbreviata y_i per denotare la soluzione numerica in x_i, cioè $y_i = y(x_i)$ (analogamente $y'_i = y'(x_i)$).

$$y_{i+1} - y_i = h y'_i$$

(*) $$y_i + 1 = y_i + h y'_i$$

Poichè dall'equazione di partenza (1*)

$$y'(x) = f(x, y(x))$$

segue

$$y'_i = f(x_i, y_i)$$

per cui da (*)

$$y_{i+1} = y_i + h f(x_i, y_i)$$

Tale metodo ha il vantaggio di essere molto semplice, tuttavia risulta essere allo stesso tempo poco accurato poichè soggetto ad un **fenomeno di propagazione dell'errore**: ad ogni passo infatti il valore y_{i+1} è affetto dall'errore dovuto all'approssimazione (2) ed è affetto dall'errore che grava su y_i, pertanto man mano che si procede nei calcoli per valutare y_1, y_2, y_3 ecc. questi errori si accumulano e si perviene a valori di y_i che possono essere molto lontani dal valore della soluzione.

Esempio di propagazione dell'errore:

Per $i=0$ il valore y_1 è affetto da errore perchè è basato sull'approssimazione (2).

$$y_1 = y_0 + hf(x_0, y_0)$$

Quando $i=1$ il valore y_2 è affetto 2 volte da errore perchè è basato sull'approssimazione (2) e perchè y_1 è a sua volta affetto da errore.

$$y_2 = y_1 + hf(x_1, y_1)$$

Quando $i=2$ il valore y_3 è affetto 3 volte da errore perchè è basato sull'approssimazione (2) e perchè y_2 è a sua volta affetto da errore.

$$y_3 = y_2 + hf(x_2, y_2)$$

1.2 Metodo di Eulero-Cauchy

Come si evince dal nome, tale metodo è una variante del metodo visto precedentemente.
A differenza della formula precedente, la formula di Eulero-Cauchy tiene conto dei due valori che sono situati attorno al punto x che si sta analizzando. Partiamo dalla relazione precedente e inseriamo i due punti x_{i-1} e x_{i+1}:

$$y'(\xi) = \frac{y_{i+1} - y_{i-1}}{2h} \simeq y'(x_i)$$

da cui

$$y_{i+1} - y_{i-1} = 2hy'_1 \Rightarrow y_{i+1} = y_{i-1} + 2hy'_i$$

ovvero:

$$y_{i+1} = y_{i-1} + 2hf(x_i, y_i)$$

Anche se più accurato del precedente tale metodo presenta lo svantaggio di non essere **autopartente**: calcolando infatti

$$y_2 = y_0 + 2hf(x_1, y_1)$$

ci troveremo nella condizione di calcolare il valore di y_1 . Da notare che tale valore non è noto a priori, tuttavia esistono alcune tecniche che permettono di approssimare y_1 mediante l'uso di derivate parziali, ma il calcolo può in alcuni casi risultare costoso oppure ci si potrebbe trovare nella situazione di non conoscere esplicitamente il valore della funzione per tale punto.

1.3 Metodi Passo-passo (Step by Step)

I metodi passo-passo permettono di trovare un'approssimazione della soluzione del problema ai valori iniziali dell'equazione differenziale di primo ordine:

$$\begin{cases} y'(x) = f(x, y(x)) \\ y(x_0) = y_0 \end{cases}$$

Applicando l'integrale ad entrambi i membri:

$$\int_{x_i}^{x_i+k} y'(x)dx = \int_{\alpha_1}^{\alpha_{i+k}} f(x, y(x))dx$$

e successivamente applicando una formula di Quadratura (Newton-Cotes):

$$y_{i+k} - y_i = h\sum_{j=0}^{k} b_j f(x_{i+j}, y_{i+j})$$

otterremo il seguente passo generico:

$$y_{i+k} = \sum_{j=0}^{k-1} \left\{ a_j y_{i+j} + h b_j f(x_{i+j}, y_{i+j} \right\} + h b_k f(x_{i+k}, y_{i+k})$$

ora:

-se $b_k = 0$ si dice che è una *formula esplicita* ovvero di tipo *predictor*

- se $b_k \neq 0$ si dice che è una *formula implicita* ovvero di tipo *corrector* **(Metodo di Milne)**

1.3.1 Metodo di Milne

Questo metodo si applica su un insieme discreto di punti utilizzando la seguente coppia di formule-differenza finite:

$$y_{i+1} = y_{i-3} + \frac{4}{3}h(2y'_{i-2} - y'_{i-1} + 2y'_i)$$
$$y_{i+1} = y_{i-1} + \frac{h}{3}(y'_{i-1} + hy'_i + y'_{i+1})$$

La prima è esplicita perchè conoscendo i valori precendenti posso calcolare y_{i+k}.
La seconda è implicita e corregge l'approssimazione dando un valore più accurato.

1.4 Metodi Runge - Kutta

Questa è una famiglia di metodi **autopartenti** e permette di calcolare in maniera esplicita i valori della y.

$$y_{i+1} = y_i + \frac{1}{2}(k_1 + k_2)$$

con

$$\begin{cases} k_1 = hf(x_i, y_i) \\ k_2 = hf(x_i + h, y_i + k_1) \end{cases}$$

Il vantaggio di questo metodo è che a differenza di Eulero-Cauchy non richiede il calcolo di derivate parziali,mentre lo svantaggio è la richiesta del calcolo di $f(x_{i+h}, y_1 + K_1)$ (nei metodi precedenti la f era calcolata in punti del tipo $f(x_i, y_i)$).

Domande Frequenti per l'esame:

Malcondizionamento
- Definizione
- Formula
- Dove intervengono le matrici mal condizionate
- Norme note

Metodo di Gauss

Fattorizzazione LU
- A che scopo è usata
- Casi applicativi
- Metodo di Choleski
  - Vantaggio

Risoluzione sistemi di equazioni lineari
- Metodo di Jacobi
  - In che caso converge
- Metodo di Gauss-Seidel
- Metodo del sovrarilassamento

Autovalori
- Se c'è un unico autovalore di valore massimo λ_1 è reale o complesso?
- Metodo delle potenze
  - A che cosa serve
  - Perchè è importante
  - Dimostrazione
- Metodo delle potenze inverse

Equazioni non lineari
- Metodo di Newton-Raphson
  - Motivazione della denominazione in metodo delle tangenti
  - Ipotesi

Interpolazione
- Tecniche di costruzione
- Matrice di Van der Monde
- Rappresentazione di Lagrange
  - Vantaggi
  - Svantaggi
- Interpolante di Lagrange
  - Come si comporta al crescere di n

Formule di quadratura
- Confronto con vantaggi e svantaggi
- Quadratura di Gauss
  - Vantaggi
  - Svantaggi
- Risoluzione con Gauss di $\int_{-1}^{1} \frac{x^5}{\sqrt{1+x^2}} dx$
- Formule di Newton-Cotes
- Formule di Gauss-Chebyshev
  - Proprietà
- Formule composite

Equazioni differenziali
- Metodo di Eulero per equazioni differeziali
  - Vantaggi
  - Svantaggi
  - Come si ricava
  - Commenti
- Confronto fra metodo di Eulero e metodo di Eulero-Cauchy
- Metodo di Runge-Kutta

Metodi impliciti ed espliciti
- Commenti
- Esempi

FFT
- Motivazioni

Conclusioni:

Queste dispense sono state fatte da studenti basandosi sugli appunti presi a lezione e sulle dispense precedentemente disponibili in rete. Per questo possono essere soggette ad errori (speriamo soltanto marginali) anche se sono state ricontrollate più volte.
Si raccomanda quindi una lettura critica e attiva che non dia nulla per scontato, che non si limiti ad imparare a memoria ciò che qui è scritto (cosa che è comunque sempre deprecabile e sconsigliata).

Detto questo, se in futuro, ad altri studenti, venga chiesto nuovamente di modificare tali dispense e quindi si presentasse il bisogno del materiale con cui è stato costruito questo pdf (appunti in .doc , .pub, ecc...) vi invitiamo a contattare gli studenti che hanno collaborato alla stesura (i contatti sono in copertina).

Buono studio a tutti!