teo rappresentazione in base dei numeri (SISTEMA che si usa per rappresentare i numeri)

Sia $\beta \in \mathbb{N}$, $\beta > 1$ β è la base di numerazione (qualsiasi base $\neq 1$)

$\forall x \in \mathbb{R}$, $x \neq 0$

$\exists!$ $p \in \mathbb{Z}$ e $\{d_i\}_{i=1}^{+\infty}$ con $d_i \in \mathbb{N}$

$0 \le d_i \le \beta - 1$, $d_1 \neq 0$ → × avere l'unicità della rappresentaz.

p → intero

$\{d_i\}$ → successione (cifre della rappresentazione)

d_i non definit. $= \beta - 1$

t.c $x = \text{sign}(x) \cdot \beta^p \sum_{i=1}^{+\infty} d_i \beta^{-i}$

sign(x): $+1$ se $x > 0$, -1 se $x < 0$; $\sum$ → MANTISSA

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

$x \in \mathbb{R} \to \tilde{x} \in F$

insieme dei numeri rappresentabili in macchina

$F(\beta, t, m, M) = \{x \in \mathbb{R} \,.\, \tilde{x} = \pm \beta^p \sum_{i=1}^{t} d_i \beta^{-i},\ d_1 \neq 0,\ 0 \le d_i \le \beta - 1\} \cup \{0\}$

β → base; t → cifre di memoria; $-m \le p \le M$ → dipende dall'hardware

$\tilde{x}$ → insieme dei n° di macchina è simmetrico

x = n° reale

$\tilde{x}$ = corrispondente n° di macchina

voglio che sia un n° di macchina perché non sia rappresentabile

$f: \mathbb{R} \to \mathbb{R}$, $f(x) \neq 0$, f razionale:

err. relativo: $\varepsilon_x = \dfrac{x - \tilde{x}}{x}$

max. ε_x: $u = \left(\frac{1}{\beta}\right)^{t-1}$ (prec. di macchina)

Se f è razionale:

$\varepsilon_{tot} = \dfrac{g(\tilde{x}) - f(x)}{f(x)}$ → funz. che voglio calcolare

(dovuto alla perturbazione dei dati iniziali) err. inerente (INEVITABILE): $\varepsilon_{in} = \dfrac{f(\tilde{x}) - f(x)}{f(x)}$

(dovuto all'algoritmo) funz. calcolata in macchina (risultato) err. algoritmico: $\varepsilon_{alg} = \dfrac{\tilde{g}(\tilde{x}) - f(\tilde{x})}{f(\tilde{x})}$ → (dovuto alle approssimazioni delle operazioni in macchina)

$\varepsilon_{in} \doteq \dfrac{f'(x)}{f(x)} \cdot x \cdot \varepsilon_x$

$\dfrac{f'(x)}{f(x)} \cdot x$ → coeff. di amplificazione

NORMA vettoriale: proprietà Sia $f: \mathbb{R}^n \to \mathbb{R}$

1) $\|x\| \ge 0$ e $\|x\| = 0 \leftrightarrow x = 0$
2) $\|\alpha x\| = \alpha \|x\|$
3) $\|x + y\| \le \|x\| + \|y\|$ (DISUG. TRIANG. ∨ CAUCHY SWARTZ)

$\|x\|_1 = \sum_{i=1}^{n} |x_i|$ $\quad \|x\|_2 = \sqrt{x^T x} = \sqrt{\sum_{i=1}^{n} x_i^2}$ $\quad \|x\|_\infty = \max_{i=1\dots n} |x_i|$

NORMA MATRICIALE: proprietà: Sia $f: \mathbb{R}^{n \times n} \to \mathbb{R}$

1) $\|A\| \ge 0$ e $(\|A\| = 0 \leftrightarrow A = 0)$
2) $\|\alpha A\| = \alpha \|A\|$
3) $\|A + B\| \le \|A\| + \|B\|$
4) $\|A \cdot B\| \le \|A\| \cdot \|B\|$

Suppongo che A sia una matrice $A \in \mathbb{R}^{n \times n}$ e che la funz. che associa x a $\|Ax\|$ sia proprietà che lega norme matriciali a norme vettoriali una funz. uniformemente continua: $x \mapsto \|Ax\|$ (Ho una norma vettoriale, una funz. continua e un insieme compatto) $X \neq \emptyset$

$\|Ax\| = \left\| A \cdot \frac{x}{\|x\|} \cdot \|x\| \right\| = \|x\| \cdot \left\| A \left(\frac{x}{\|x\|} \right) \right\| \le \|x\| \cdot \|A\|$

$\|A\|$ è il max perché → versore

definisco $\|A\| = \max_{x \in S} \|Ax\|$ sull'insieme $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$, *

S è chiuso e limitato in $\mathbb{R}^n$ → è compatto di $\mathbb{R}^n$

2ª prop. delle norme

→ $\|Ax\|_v \le \|A\|_m \cdot \|x\|_v$ (norma matriciale; norme vettoriali)

norma matriciale indotta dalla norma vettoriale

*il max di una funz. continua su un insieme chiuso e limitato ∃ per teo WEIERSTRASS

$\|A\|_1 = \max_{j=1\dots n} \sum_{i=1}^{m} |a_{ij}|$ $\quad \|A\|_\infty = \max_{i=1\dots n} \sum_{j=1}^{m} |a_{ij}|$ $\quad \|A\|_2 = \sqrt{\varphi(A^T A)}$ dove $\varphi(A) = \max_{i=1\dots n} |\lambda_i|$

(dove λ è lo scalare (anche complesso) t.c $Ax = \lambda x$ con $x \neq \emptyset$ autovettore) — modulo dell'autovalore di modulo max

n° di condizionamento del sist. lineare:

$1 = \|A \cdot A^{-1}\| = \|I\| \leq \|A\| \cdot \|A^{-1}\| \quad \rightarrow \quad \|A\| \cdot \|A^{-1}\| \geq 1$

(4ª prop. delle norme)

$\|I\| = \max \|Ix\| = \max_{\|x\|=1} \|x\| = 1$ $\quad \{\|x\| = 1\}$

(per una qualsiasi norma matriciale indotta, la norma della matrice identica è 1)

sistema lineare: $Ax = b$

INPUT: $A \in \mathbb{R}^{n \times n}, \mathbb{C}^{n \times n}$ (matrice quadrata); $b \in \mathbb{R}^n, \mathbb{C}^n$;

OUTPUT: $x \in \mathbb{R}^n, \mathbb{C}^n$

hp: A è invertibile $\rightarrow \exists B$ t.c. $AB = B \cdot A = I \rightarrow B = A^{-1}$

$\rightarrow \exists!$ la soluz. del sistema $\rightarrow x = A^{-1} \cdot b$

studio del condizionamento: quanto è distante la soluzione del problema perturbato ($\hat{x}$) dalla soluz. del problema iniziale (x)? Per semplicità perturbiamo solo il termine noto b.

$Ax = b$

sistema perturbato (risolto in macchina)

$A\hat{x} = b + \delta b \rightarrow \hat{x} = A^{-1}b + A^{-1}\delta b$ (perturbazione)

risolvo il sist. sui dati perturbati:

$$\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \rightsquigarrow \begin{bmatrix} \tilde{b_1} \\ \vdots \\ \tilde{b_n} \end{bmatrix} = \begin{bmatrix} b_1(1+\varepsilon_1) \\ \vdots \\ b_n(1+\varepsilon_n) \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} + \underbrace{\begin{bmatrix} b_1\varepsilon_1 \\ \vdots \\ b_n\varepsilon_n \end{bmatrix}}_{=\delta b}$$

dist $(\hat{x}, x)$:

$$\varepsilon_{relativo} = \frac{\|\hat{x} - x\|}{\|x\|} = \frac{\|\cancel{A^{-1}b} + A^{-1}\delta b - \cancel{A^{-1}b}\|}{\|x\|} = \frac{\|A^{-1}\delta b\|}{\|x\|} \leq \frac{\|A^{-1}\| \cdot \|\delta b\|}{\|x\|} \leq \frac{\|A^{-1}\| \cdot \|\delta b\|}{\|b\|} \cdot \|A\|$$

(norma matriciale indotta dalla norma vettoriale)

$\|Ax\| = \|b\| \leq \|A\| \cdot \|x\| \rightarrow \|x\| \geq \frac{\|b\|}{\|A\|}$ (↓ norma matric. indotta dalla vettoriale)

per fare una maggiorazione il denominatore deve essere minore

$\rightarrow \frac{\|\hat{x} - x\|}{\|x\|} \leq \|A^{-1}\| \cdot \|A\| \cdot \frac{\|\delta b\|}{\|b\|}$

$\|A^{-1}\| \cdot \|A\|$: n° di condizionamento (se è grande → sist. mal condizionato; se è piccolo → sist. ben condizionato)

$\frac{\|\delta b\|}{\|b\|}$: errore relativo sui dati iniziali in norma

($\|\hat{x} - x\|$)

(Studiare la stabilità dell'algoritmo è un problema di analisi del condizionamento)

teo: $\varepsilon_{TOT} \doteq \varepsilon_{in} + \varepsilon_{alg}$ (analisi al primo ordine dell'errore) (vale se g è una funz. razionale)

Sia $g: \mathbb{R} \rightarrow \mathbb{R}$: $\varepsilon_{TOT} = \frac{g(\tilde{x}) - g(x)}{g(x)} = \frac{g(\tilde{x}) - g(x) + g(\tilde{x}) - g(\tilde{x})}{g(x)} = \frac{g(\tilde{x}) - g(\tilde{x})}{g(x)} + \frac{g(\tilde{x}) - g(x)}{g(x)} =$ (ε_{in})

$= \frac{g(\tilde{x}) - g(\tilde{x})}{g(x)} \cdot \frac{g(\tilde{x})}{g(\tilde{x})} + \varepsilon_{in} = \varepsilon_{alg} \cdot \left(\frac{g(\tilde{x}) + g(x) - g(x)}{g(x)}\right) + \varepsilon_{in} =$ (ε_{alg})

$= \varepsilon_{alg}\left(\frac{g(\tilde{x}) - g(x)}{g(x)} + \frac{g(x)}{g(x)}\right) + \varepsilon_{in} = \varepsilon_{alg}(\varepsilon_{in} + 1) + \varepsilon_{in} \doteq \varepsilon_{alg} + \varepsilon_{in}$

(ε_{alg} ci dice la STABILITÀ; ε_{in} ci dice il condizionamento)

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

FATTORIZZAZIONE LU = prodotto di 2 matrici triangolari

$A \in \mathbb{R}^{n\times n}$ è fattorizzabile nella forma LU se

$\exists \; L \in \mathbb{R}^{n\times n}$ matrice triang. inferiore con $l_{ii} = 1 \quad (i = 1 \dots n)$

e $U \in \mathbb{R}^{n\times n}$ = triang. superiore t.c. $A = L \cdot U$

$Ax = b \longleftrightarrow LUx = b \longleftrightarrow \begin{cases} Ly = b \rightarrow \text{SOST. in avanti} \\ Ux = y \rightarrow \text{sost. all'indietro} \end{cases}$

(posso risolvere il sist. lineare $Ax = b$ risolvendo 2 sist. triangolari se ho una fattorizzazione LU)

condiz. sufficiente per la fattorizzazione LU:

teo: Sia $A \in \mathbb{R}^{n\times n}$: se le sottomatrici principali di testa A_k (di ordine k) sn invertibili per $k = 1 \dots n-1 \rightarrow \exists!$ la fattorizzazione LU di A.

DIMO per induzione su n (dimensione della matrice A):

dimostro per le matrici di ordine 1, assumo che il teo sia vero fino alle matrici di ordine n-1, dimostro per n.

PASSO BASE: n = 1 (un solo elemento) $(A \in \mathbb{R}^{1\times 1})$ $a = L \cdot U$ (unica fattorizzazione) $\rightarrow a = 1 \cdot u \rightarrow a = u$

ORDINE di una matrice = DIMENSIONE = n° di righe e di colonne

$\forall$ ipotesi del teorema per n = 1 XK non si hanno sottomatrici (il teo. non pone condizioni)

$\exists!$ fatt. LU per n = 1

Suppongo che il teo. sia vero fino all'ordine n-1 e ora lo dimostro per il passo successivo n: (partizionamento a blocchi)

scompongo A: cerco una fatt LU t.c. $A = LU$

$$\left[\begin{array}{c|c} A_{n-1} & z \\ \hline w^T & \alpha \end{array}\right] = \left[\begin{array}{c|c} L_{n-1} & 0 \\ \hline x^T & 1 \end{array}\right] \cdot \left[\begin{array}{c|c} U_{n-1} & y \\ \hline 0 & \gamma \end{array}\right]$$

(A_{n-1}: matrice; z: vettore colonna; w^T: vettore riga; α: scalare)

qst termini sn incognite → voglio determinarli in modo univoco e vedere se sn unici

PRODOTTO RIGA × COLONNA:

$$\leftrightarrow \begin{cases} A_{n-1} = L_{n-1} \cdot U_{n-1} \\ z = L_{n-1} \cdot y \\ w^T = x^T \cdot U_{n-1} \\ \alpha = x^T y + \gamma \end{cases}$$

(la MOLTIPLIC. tra matrici non gode di prop. commutativa)

ho 2 hp:

- INDUTTIVA: $\exists!$ fattoriss. LU delle matrici di ordine inferiore a n (da ordine 1 a n-1) se queste sn invertibili
- del teorema: dice che le sottomatrici princ. di testa di A per $k = 1 : n-1$ sn invertibili

• $A_{n-1} = L_{n-1} \cdot U_{n-1}$: sottomatrici princip. di testa di A_{n-1} coincidono con quelle di A (A_{n-1} è una sottomatrice di A)

hp del teorema: le sottomatrici princ. di testa di A sn invertibili (fino all'ordine n-2)

hp induttiva: $\exists!$ fatt. LU di A_{n-1}

→ L_{n-1} e U_{n-1} sn univocamente determinati e sn i fattori dell'unica fatt. LU di A_{n-1}

• $z = L_{n-1} \cdot y$: A_{n-1} è invertibile per l'hp del teorema (per il teo di BINET)

$\det(A_{n-1}) = \det(L_{n-1} \cdot U_{n-1}) \rightarrow \det(A_{n-1}) = \det(L_{n-1}) \cdot \det(U_{n-1})$

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

$\to \det(A_{n-1}) \neq 0$ ($\times K$ invertibile) $\to \underbrace{\det(L_{n-1})}_{=1} \cdot \det(U_{n-1}) \neq 0$

$\to L_{n-1}$ e U_{n-1} sn invertibili

$\to$ l'equazione è univocamente determinata : $\boxed{y = (L_{n-1})^{-1} z}$

are
Rari

$w^T = x^T \cdot U_{n-1}$; $\boxed{x^T = w^T (U_{n-1})^{-1}}$ $\quad \times K (U_{n-1})$ è invertibile

$\alpha = x^T \cdot y + \gamma$: $\boxed{\gamma = \alpha - x^T y}$

↓ ↙

sn conosciuti

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

li

$\to$ gli elementi della fattorizzazione sn univocamente determinati

$\to \exists!$ la fatt. LU di A.

---

li

Teorema di GERSHGORIN:

Sia $A \in \mathbb{C}^{n \times n}$, se λ è autovalore di $A \to \lambda \in \bigcup_{i=1}^{n} K_i$ dove

$$K_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \right\} \qquad i = 1 \dots n$$

centri ↓ raggi ↓

DIMO: Parto dalla definizione di autovalore e autovettore:

λ è autovalore di A se $\exists x \neq \emptyset$ per cui $Ax = \lambda x$, con x autovettore associato all'autovalore λ.

posso scrivere: (PROD. RIGA COLONNA)

$$\sum_{j=1}^{n} a_{ij} x_j = \lambda x_i \qquad \text{per } i = 1 \dots n$$

(prodotto della i-esima riga per il vettore x = i-esima componente del vettore λx (vale $\forall$ componente))

porto a dx il caso per $j = i$: (componente diagonale)

$$\sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j = \lambda x_i - a_{ii} x_i \qquad i = 1 \dots n$$

$$\downarrow$$

$$\sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j = (\lambda - a_{ii}) x_i \qquad \to \text{passo ai valori assoluti:}$$

* se fossero componenti tutte uguali $\to \lambda \in \Lambda$ cerchi di modulo max
· la scelta del cerchio dipende dalla conoscenza dell'autovettore

$$\left| \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j \right| = \left| (\lambda - a_{ii}) x_i \right| = |\lambda - a_{ii}| |x_i| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \cdot |x_j| \qquad \text{per } i = 1 \dots n$$

(poiché $|x \cdot y| = |x| \cdot |y|$ (modulo prodotto = prod. dei moduli)) (poiché $|x+y| \leq |x| + |y|$ (modulo somma ≤ somma dei moduli))

$$|\lambda - a_{ii}| |x_i| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| |x_j| \qquad i = 1 \dots n$$

Scelgo k t.c. $|x_k| = \max_{j=1 \dots n} |x_j|$ dove K = indice della componente di modulo max tra le componenti dell'autovettore x relativo all'autovalore λ

$\to x_k$ è univocamente determinato $\to x_k \neq \emptyset$ $\to$ posso dividere per $|x_k|$

sostituisco $i = k$: ($x \neq 0$ $\forall k$ è autovettore)

considero il modulo $\forall k$ x_k potrebbe essere complesso

$$|\lambda - a_{kk}| \frac{|x_k|}{|x_k|} \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}| \frac{|x_j|}{|x_k|} \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}|$$

(al max: $j = k$) $\quad \leq 1$ $\forall k$ $|x_k|$ è il max e la disequazione mantiene lo stesso verso $\forall k$ $x_k \neq \emptyset$ *

$$\to |\lambda - a_{kk}| \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}| \to \lambda \in K_k \to \lambda \in \bigcup_{i=1}^{n} K_i$$

Teo: Sia $A \in \mathbb{R}^{n\times n}$ e $\det(A_k) \neq 0$ per $1 \le k \le p \le n-1$
supponga che le prime p sottomatrici siano invertibili
→ i primi p passi del procesa di fatt. LU si possono eseguire e $a_{kk}^{(k-1)} \neq 0$ per $1 \le k \le p \le n-1$

risoluzione iterativa di sist. lineari $Ax=b$
successione di vettori $x^{(k)}$ che sotto opportune condizioni convergono alla soluz. del sist. lineare:

$$\left(\{x^{(k)}\}_{k\in\mathbb{N}} = x^{(k)} \in \mathbb{R}^n = x^{(k)} \underset{k\to\infty}{\to} x \right)$$

convergenza di vettori: $x^{(k)} \to x \leftrightarrow \lim_{k\to\infty} \|x^{(k)} - x\| = 0$

(la convergenza non dipende dalla norma scelta)

$$\|x^{(k)} - x\|_\infty = \max_{i=1\dots n} |x_i^{(k)} - x_i|$$

$$\to 0 \le |x_i^{(k)} - x_i| \le \|x^{(k)} - x\|_\infty \quad i = 1\dots n$$

$\downarrow 0$ per teo carabinieri $\qquad \downarrow 0$

$$\left[x_i^{(k)} \to x_i \right]$$

MASTER COPY
Tel. 050 8312126
Cell. 388 9837745

$Ax=b$. Decomposo A ~ differenza di due matrici: $A = M - N$ e assumo che M sia invertibile

$$Ax=b \leftrightarrow (M-N)x = b \leftrightarrow Mx - Nx = b \leftrightarrow Mx = Nx + b$$

x è soluzione se e solo se soddisfa:

$$x = \underbrace{M^{-1}\cdot N}_{P} x + \underbrace{M^{-1}\cdot b}_{q}$$

P = (matrice di iterazione)

(La condiz. di applicabilità per i metodi iterativi è che M sia invertibile)

chiamo $g(x)$ la funzione: $g(x) = Px + q = M^{-1}Nx + M^{-1}b$
dove $g(x)$ è una funzione: $g: \mathbb{R}^n \to \mathbb{R}^n$ - $g(x)$ è continua.

Considero qst procedimento per generare la successione $\begin{cases} x^{(0)} \in \mathbb{R}^n \\ x^{(k+1)} = g(x^{(k)}) \quad k \ge 0 \end{cases}$
· scelgo un vettore iniziale
· genero la successione (una sequenza di vettori)

Teo: Se $\lim_{k\to+\infty} x^{(k)} = \hat{x} \in \mathbb{R}^n \to A\hat{x} = b$ (se la successione converge, $\hat{x}$ è la soluz. del sistema lineare)

DIMO: per costruzione: $x^{(k+1)} = g(x^{(k)}) = Px^{(k)} + q$
faccio il lim per $k\to+\infty$: $\quad \hat{x} = g(\hat{x}) = P\hat{x} + q = M^{-1}\cdot N\hat{x} + M^{-1}\cdot b$ (* perché $g(x)$ è continua)

moltiplico per M: $M\hat{x} = N\hat{x} + b$
$M\hat{x} - N\hat{x} = b$
$(M-N)\hat{x} = b$
$A\hat{x} = b$

---

definizione: Il met. iterativo $x^{(k+1)} = g(x^{(k)}) = M^{-1}\cdot Nx^{(k)} + M^{-1}\cdot b$ con $(M-N)=A$ è convergente se la successione generata $\begin{cases} x^{(0)} \in \mathbb{R}^n \\ x^{(k+1)} = g(x^{(k)}) \end{cases}$ converge $\forall x^{(0)} \in \mathbb{R}^n$
($\forall$ scelta del punto iniziale)