1 Lezione del 31-03-25

Riprendiamo la trattazione dei metodi iterativi per i sistemi lineari.

1.0.1 Criteri di convergenza di matrici di iterazione

Avevamo visto che un problema perchè un metodo di punto sia convergente è che la sua matrice H in:

$$x = Hx + c$$

sia convergente, cioè avesse norma spettrale $\rho(H) < 1$.

Vediamo che vale il seguente problema.

Teorema 1.1: Criteri di convergenza di matrici di iterazione

Se $A \in \mathbb{C}^{n \times n}$ rispetta una delle seguenti condizioni:

1. A è ha predominanza diagonale forte;
2. A è irriducibile e a predominanza diagonale debole;

allora sia il metodo di Jacobi che il metodo di Gauss-Seidel danno iterazioni convergenti applicati ad un sistema $Ax = b$.

1. La dimostrazione del teorema comincia dall'imporre che $a_{ii} \neq 0$, $\forall i = 1, ..., n$. Avevamo visto che questa condizione può essere forzata scambiando A con una qualche permutazione ΠA.

   Abbiamo allora che la condizione è ovvia nel caso di predominanza diagonale forte (1), mentre nel caso di matrice irriducibile a predominanza diagonale debole (2), se esistesse un $a_{ii} = 0$ per qualche i allora applicando la predominanza diagonale debole si avrebbe $a_{ij} = 0$ $(0 \geq \sum_{i \neq j}^{n} a_{ij})$, che sarebbe assurdo (la matrice non può essere singolare). Notiamo che in questo caso nemmeno una matrice di permutazione Π potrebbe portarci a diagonale non negativa.

2. Dimostriamo la convergenza per i due metodi:

   - **Jacobi:** si ha che:

     $$H_J = D^{-1}(E+F) = \begin{pmatrix} 0 & ... & -\frac{a_{ij}}{a_{ii}} \\ ... & 0 & ... \\ -\frac{a_{ij}}{a_{ii}} & ... & 0 \end{pmatrix}$$

     e allora la norma a infinito di H_J sarà:

     $$|H_J|_\infty = \max_{i=1,...,n} \sum_{i \neq j}^{n} \left| \frac{a_{ij}}{a_{ii}} \right|$$

     Se A è a predominanza diagonale forte (1), allora $|a_{ii}| > \sum_{j \neq i}^{n} |a_{ij}|$, e quindi $|H_J|_\infty < 1$ e $\rho(H_j) < 1$.
     Se A è invece irriducibile a predominanza diagonale debole (2), sarà che nuovamente $|H_J|_\infty \leq 1$, quindi gli autovalori di A sono $|\lambda| \leq 1$, ma non esiste $|\lambda| = 1$ in quanto starebbe sul bordo di un cerchio di Gershgorin, quindi di tutti, ma esiste almeno un cerchio $\mathcal{F}_i(A)$ con raggio $r < 1$.

- **Gauss-Seidel:** supponiamo che esista λ^* autovalore di H_{GS}:

  $$H_{GS} = (D - E)^{-1} F$$

  tale che $|\lambda^*| \geq 1$. Allora dovrà valere:

  $$0 = \det(\lambda^* I - H_{GS}) = \det\left(\lambda^* I - (D - E)^{-1} F\right) = \det\left((D - E)^{-1} \left(\lambda(D - E) - F\right)\right)$$

  che per Binet-Cauchy dà:

  $$0 = \det\left((D - E)^{-1}\right) \cdot \det\left(\lambda^*(D - E) - F\right)$$

  Il primo termine è chiaramente $\neq 0$, e quindi si considera:

  $$0 = \det\left(\lambda^*(D - E) - F\right) = (\lambda^*)^n \det\left(D - E - (\lambda^*)^{-1} F\right)$$

  cioè:

  $$0 = \det\left(D - E - (\lambda^*)^{-1} F\right)$$

La matrice che troviamo è quella dove gli elementi sopra la diagonale hanno modulo $\leq$ ai corrispondenti nella matrice A. Se A è dominante diagonale forte o dominante diagonale debole, quindi, conserva tale proprietà, e conserva anche l'irriducibilità. Quindi, necessariamente, è non singolare (altrimenti anche A sarebbe singolare), e l'ipotesi $|\lambda^*| \geq 1$ è assurda, cioè $|\lambda^*| < 1$.

$\square$

1.1 Sistemi rettangolari

Vediamo quindi come risolvere sistemi **rettangolari**, cioè del tipo $Ax = b$ con $A \in \mathbb{C}^{m \times n}$ rettangolare. In questo caso si avranno m equazioni in n incognite, quindi con $x \in \mathbb{C}^n$ e $b \in \mathbb{C}^m$. Supponiamo poi che A sia di rango massimo, ovvero $\operatorname{rank}(A) = \min(m, n)$.

Scopriamo che anche in questo scenario si può applicare l'eliminazione di Gauss, e nel caso $m < n$, si può immaginare di rendere il sistema quadrato aggiungendo $n - m$ equazioni fittizie della forma $0 = 0$, lasciando $n - m$ parametri liberi. Diciamo che questo tipo di sistemi si dicono **sottodeterminati**.

Caso più interessante sarà quello dei sistemi **sovradeterminati**, cioè dove $n < m$ e si hanno più equazioni che incognite. In questo caso l'applicazione dell'eliminazione di Gauss porta ad n equazioni determinate, e $m - n$ equazioni della forma $0 = ?$ dove $?$ sono i termini che si formano ai termini noti nelle ultime $m - n$ righe. Se esiste almeno un termine noto $? \neq 0$, allora il sistema è impossibile (ed è questo il caso tipico). Questa affermazione è essenzialmente il teorema di Rouchè-Capelli, dove si impone:

$$\operatorname{rank}(A|b) = \operatorname{rank}(A)$$

Il problema che ci poniamo in questi casi è quello di minimizzare il residuo, cioè posto:

$$x - Ab = r(x)$$

cercare di risolvere il problema di ottimizzazione:

$$r^* = \min_{x \in \mathbb{C}^n} |b - Ax|_2 = \min_{x \in \mathbb{C}^n} |r(x)|_2$$

Questo tipo di problema viene detto **problema dei minimi quadrati**. Osserviamo che il problema è equivalente a risolvere:

$$(r^*)^2 = \min_{x \in \mathbb{C}^n} |b - Ax|_2^2 = \min_{x \in \mathbb{C}^n} |r(x)|_2^2$$

Cioè ci possiamo liberare della radice della norma quadratica.

Potremo innanzitutto chiederci come trovare il punto x^* che dà r^*, o in principio se questo problema ammette soluzione. Facciamo allora un esempio del caso $x \in \mathbb{R}^n$ Considereremo la funzione:

$$\psi(x) = |b - Ax|_2^2 = (b - Ax)^T (b - Ax) : \mathbb{R}^n \to \mathbb{R}$$

questa funzione è convessa, e quindi il punto x^* è quello che soddisfa $\nabla \psi(x) = 0$ (gradiente nullo). Vediamo allora le derivate parziali, osservando che:

$$|b - Ax|_2^2 = |r(x)|_2^2 = \sum_{i=1}^{m} r_i(x)^2$$

e:

$$r_i(x)^2 = \left(b_i - \sum_{j}^{n} a_{ij} x_j \right)^2$$

Calcoliamo quindi la derivata parziale vera e propria rispetto alla s-esima componente:

$$\frac{\partial}{\partial x_s} \left(r_i(x)^2 \right) = 2 \left(b_i - \sum_{j=1}^{n} a_{ij} x_j \right) (-a_{is})$$

da cui:

$$\frac{\partial}{\partial x_s} (\psi(x)) = \sum_{i=1}^{n} \frac{\partial}{\partial x_s} \left(r_i(x)^2 \right) = 2 \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{is} a_{ij} x_j - a_{is} b_i \right)$$

$$= 2 \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{is} a_{ij} x_j - \sum_{i=1}^{n} a_{is} b_i \right) = 2 \left(\sum_{i=1}^{m} a_{is} \sum_{j=1}^{n} a_{ij} x_j - \sum_{i=1}^{m} a_{is} b_i \right)$$

$$= 2 \left(\left(A^T A x \right)_s - \left(A^T b \right)_s \right) = 2 \left(A^T A x - A^T b \right)_s$$

Si ha quindi che:

$$\nabla \psi(x) = 0 \Leftrightarrow A^T A x - A^T b = 0 \Leftrightarrow A^T A x = A^t b$$