



## RISOLUZIONE DI SISTEMI LINEARI: METODI ITERATIVI (Metodi di Richardson)

In questa lezione proseguiamo la discussione sui metodi iterativi per la risoluzione di sistemi lineari, concentrandoci in particolare sui cosiddetti metodi di tipo gradiente. I metodi di tipo gradiente sfruttano l'informazione relativa al gradiente di una funzione opportunamente associata al sistema lineare. Questi metodi rappresentano una generalizzazione di quelli esaminati fino ad ora, ossia i metodi di Jacobi, di Gauss-Seidel (GS) e del rilassamento (o SOR).

Iniziamo la trattazione sui metodi del gradiente con l'introduzione del cosiddetto metodo di Richardson, che si configura come una generalizzazione degli schemi iterativi che abbiamo già visto.

Nella precedente lezione, abbiamo visto che un metodo iterativo generico può essere formulato tramite uno schema del tipo

$$P(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = \mathbf{r}^{(k)},$$

dove  $P$  è la matrice che abbiamo chiamato di preconditionamento ed  $\mathbf{r}^{(k)}$  il residuo. Il metodo di Richardson introduce un parametro aggiuntivo che chiameremo  $\alpha$  che va a moltiplicare il residuo  $\mathbf{r}^{(k)}$

$$P(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = \alpha \mathbf{r}^{(k)}.$$

Pertanto, invece di considerare il residuo al passo  $k$ , moltiplicheremo il residuo per un parametro  $\alpha$ . Questo parametro  $\alpha$  funge da parametro di accelerazione, in modo analogo a quanto osservato con il parametro  $\omega$  introdotto per il metodo di rilassamento SOR.

La matrice di iterazione del metodo di Richardson risulta essere

$$R_\alpha = I - \alpha P^{-1}A,$$

da cui ponendo  $\alpha = 1$ , ritroviamo la matrice di iterazione  $B = I - P^{-1}A$ .

Il metodo di Richardson può essere considerato come un'estensione dei metodi iterativi visti finora, introducendo un parametro di accelerazione  $\alpha$  che serve a ottenere una convergenza più rapida possibile.

La prima domanda che ci poniamo è come determinare questo parametro di accelerazione  $\alpha$ . A tal fine, occorre studiare l'analisi di convergenza del metodo di Richardson, dalla quale possiamo derivare come la matrice di iterazione dipenda da  $\alpha$ . Successivamente, dovremo comprendere come scegliere  $\alpha$  in modo tale che la matrice di iterazione sia, innanzitutto, convergente, ovvero abbia un raggio spettrale minore di uno, e, se possibile, che il raggio spettrale sia il più piccolo possibile, ottimizzando così la velocità di convergenza del metodo.

Osserviamo innanzitutto che gli autovalori di  $R_\alpha = I - \alpha P^{-1}A$ , sono

$$\lambda_k = 1 - \alpha \mu_k \quad k = 1, \dots, n,$$

dove  $\mu_k$  è il generico autovalore di  $P^{-1}A$ . Quindi  $\mu_k$  è l'autovalore della matrice  $A$  preconditionata (ovvero la matrice  $A$  dopo l'applicazione della matrice inversa di  $P$ ). Dunque, per avere convergenza del metodo di Richardson dovrà necessariamente essere che ogni autovalore  $\lambda_k$  abbia un modulo minore di uno

$$|\lambda_k| = |1 - \alpha \mu_k| < 1 \quad k = 1, \dots, n.$$

Vediamo adesso come questa relazione, che è una relazione di convergenza, si caratterizza rispetto ad  $\alpha$  e quindi come dipende da  $\alpha$ .



Nell'ipotesi in cui gli autovalori  $\mu_k \in \mathbb{R}^+$ , ovvero  $\mu_k$  sono numeri reali positivi, si può verificare che il limite di convergenza è  $\alpha = \frac{2}{\mu_{\max}}$ , il quale corrisponde ad avere raggio spettrale  $\rho = 1$ . Di conseguenza, il metodo di Richardson converge se e solo se

$$0 < \alpha < \frac{2}{\mu_{\max}},$$

dove  $\mu_{\max} = \max_k \{\mu_k(P^{-1}A)\}$ , ovvero il massimo autovalore della matrice  $P^{-1}A$ . Inoltre, se  $A$  e  $P$  sono matrici simmetriche e definite positive (SDP) allora sicuramente  $\mu_k \in \mathbb{R}^+$ , ovvero  $\mu_k$  sono numeri reali positivi. In particolare, anche il massimo  $\mu_{\max}$  e il minimo  $\mu_{\min}$  autovalore di  $P^{-1}A$  saranno positivi  $\mu_{\max} > \mu_{\min} > 0$ .

Nel caso di  $A$  e  $P$  matrici SDP (e quindi anche  $P^{-1}A$  matrice SDP) si può dimostrare che la condizione per la quale si ottiene il valore ottimale  $\alpha_{\text{ott}}$ , ovvero il valore per cui il raggio spettrale della matrice di iterazione  $R_\alpha$  è minimo (ossia il valore che minimizza il raggio spettrale), corrisponde alla soluzione del problema

$$1 - \alpha_{\text{ott}}\mu_{\min} = -1 + \alpha_{\text{ott}}\mu_{\max},$$

da cui si trova

$$\alpha_{\text{ott}} = \frac{2}{\mu_{\min} + \mu_{\max}}.$$

Quindi esiste in questo caso una relazione precisa per il parametro di accelerazione ottimale in funzione degli autovalori minimo e massimo di  $P^{-1}A$ . In corrispondenza del valore ottimale  $\alpha_{\text{ott}}$ , è possibile determinare il raggio spettrale ottimale  $\rho_{\text{ott}}(R_\alpha)$ . Infatti, il calcolo si riconduce alla risoluzione di dell'equazione

$$\rho_{\text{ott}} = 1 - \alpha_{\text{ott}}\mu_{\min},$$

da cui andando a sostituire il valore  $\alpha_{\text{ott}}$  si ricava

$$\rho_{\text{ott}} = 1 - \alpha_{\text{ott}}\mu_{\min} = 1 - \frac{2\mu_{\min}}{\mu_{\min} + \mu_{\max}} = \frac{\mu_{\max} - \mu_{\min}}{\mu_{\min} + \mu_{\max}} = \left(\frac{\mu_{\max}}{\mu_{\min}} - 1\right) / \left(\frac{\mu_{\max}}{\mu_{\min}} + 1\right),$$

dove nell'ultimo passaggio abbiamo diviso numeratore e denominatore per  $\mu_{\min}$ . Ricordando che  $\mu_{\max}$  e  $\mu_{\min}$  sono il massimo e il minimo autovalore di  $P^{-1}A$  abbiamo che il condizionamento in norma due di  $P^{-1}A$  è esattamente il rapporto tra il massimo e il minimo autovalore (dato che stiamo supponendo che  $P^{-1}A$  sia SDP)

$$K_2(P^{-1}A) = \frac{\mu_{\max}}{\mu_{\min}}$$

Con questa ultima caratterizzazione, e dall'ultima formula ricavata per  $\rho_{\text{ott}}$ , si ottiene che

$$\rho_{\text{ott}}(R_\alpha) = \frac{K_2(P^{-1}A) - 1}{K_2(P^{-1}A) + 1}.$$

Quindi il raggio spettrale ottimale è esattamente espresso in funzione del numero di condizionamento di  $P^{-1}A$ .

In sintesi, abbiamo mostrato come scegliere  $\alpha$  per assicurare la convergenza per metodo di Richardson (con parametro  $\alpha$  costante) nel caso  $P^{-1}A$  è SDP, ovvero

$$0 < \alpha < \frac{2}{\mu_{\max}},$$

e l' $\alpha$  ottimale è

$$\alpha_{\text{ott}} = \frac{2}{\mu_{\min} + \mu_{\max}},$$



ed è quello in corrispondenza del quale abbiamo una velocità di convergenza ottimale quindi un raggio spettrale della matrice di interazione il più piccolo possibile

$$\rho_{ott}(R_\alpha) = \frac{\left(\frac{\mu_{max}}{\mu_{min}} - 1\right)}{\left(\frac{\mu_{max}}{\mu_{min}} + 1\right)} = \frac{K_2(P^{-1}A) - 1}{K_2(P^{-1}A) + 1}.$$

Esaminiamo ora come generalizzare ulteriormente questo approccio, introducendo una scelta dinamica del parametro  $\alpha$ . La necessità di un parametro dinamico nasce dal fatto che, osservando la relazione

$$\alpha_{ott} = \frac{2}{\mu_{min} + \mu_{max}},$$

risulta evidente che la determinazione esatta di  $\alpha_{ott}$  richiede la conoscenza degli autovalori minimo e massimo della matrice. Tuttavia, tale informazione non è sempre disponibile. Di conseguenza, è opportuno individuare strategie alternative che consentano di determinare  $\alpha$  in modo automatico, senza la necessità di conoscere a priori il suo valore ottimale.

Per questo motivo, introduciamo i metodi di Richardson dinamici, nei quali il parametro  $\alpha$  può essere aggiornato a ogni iterazione in base a un criterio adattivo. Questa classe di metodi conduce in particolare al metodo del gradiente.

Analizziamo dunque il metodo del gradiente nella sua formulazione generale. L'idea di base consiste nel selezionare, a ogni passo dell'iterazione di Richardson, un parametro di accelerazione variabile determinato da un criterio automatico. L'obiettivo resta quello di massimizzare la velocità di convergenza del metodo. In altri termini, intendiamo ottenere lo stesso vantaggio del metodo di Richardson con parametro costante - ovvero garantire una convergenza ottimale - ma senza dover dipendere dalla conoscenza degli autovalori della matrice.

Iniziamo con una riscrittura del metodo di Richardson

$$P(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = \alpha \mathbf{r}^{(k)}.$$

Osserviamo che, se applichiamo  $P^{-1}$  a sinistra e a destra troviamo

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha P^{-1} \mathbf{r}^{(k)} = \mathbf{x}^{(k)} + \alpha \mathbf{z}^{(k)},$$

dove, nell'ultimo passaggio, abbiamo introdotto il vettore

$$\mathbf{z}^{(k)} = P^{-1} \mathbf{r}^{(k)}.$$

Il vettore  $\mathbf{z}^{(k)}$  è chiamato residuo preconditionato, dato che  $P$  è il preconditionatore. Abbiamo quindi trovato una relazione ricorsiva per  $\mathbf{x}^{(k)}$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha \mathbf{z}^{(k)}.$$

In maniera analoga possiamo trovare una relazione ricorsiva per  $\mathbf{r}^{(k)}$ . Più precisamente possiamo scrivere che il residuo al passo  $k + 1$  è

$$\mathbf{r}^{(k+1)} = \mathbf{b} - A\mathbf{x}^{(k+1)},$$

e tenendo conto che  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha \mathbf{z}^{(k)}$ , abbiamo

$$\mathbf{r}^{(k+1)} = \mathbf{b} - A\mathbf{x}^{(k)} - \alpha A\mathbf{z}^{(k)} = \mathbf{r}^{(k)} - \alpha A\mathbf{z}^{(k)},$$

da cui

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha A\mathbf{z}^{(k)},$$

Osserviamo che entrambe le successioni  $\mathbf{x}^{(k)}$  ed  $\mathbf{r}^{(k)}$  possono essere viste in maniera ricorsiva attraverso il vettore  $\mathbf{z}^{(k)}$



$$\begin{aligned} \mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} + \alpha \mathbf{z}^{(k)}, \\ \mathbf{r}^{(k+1)} &= \mathbf{r}^{(k)} - \alpha A \mathbf{z}^{(k)}, \end{aligned}$$

il quale altro non è che un residuo preconditionato attraverso la matrice  $P^{-1}$ .

Desideriamo ora sfruttare queste informazioni per riscrivere il metodo di Richardson in una forma più interessante dal punto di vista computazionale. Se riassumiamo le informazioni precedenti abbiamo

$$\begin{cases} P \mathbf{z}^{(k)} = \mathbf{r}^{(k)} \\ \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha \mathbf{z}^{(k)}, \\ \mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha A \mathbf{z}^{(k)} \end{cases}$$

dove abbiamo espresso  $\mathbf{z}^{(k)} = P^{-1} \mathbf{r}^{(k)}$  come soluzione di un sistema lineare  $P \mathbf{z}^{(k)} = \mathbf{r}^{(k)}$  la cui matrice dei coefficienti è  $P$  e il cui termine noto è  $\mathbf{r}^{(k)}$ . Questo che abbiamo appena scritto è il metodo Richardson con parametro  $\alpha$  con parametro fissato (o costante) per ogni iterazione  $k$ .

Introduciamo adesso un nuovo metodo di Richardson con parametro  $\alpha$  dinamico. Per ottenere una scelta dinamica, è sufficiente sostituire ad  $\alpha$  il parametro  $\alpha_k$ , dove  $\alpha_k$  è un coefficiente che varia a ogni iterazione

$$\begin{cases} P \mathbf{z}^{(k)} = \mathbf{r}^{(k)} \\ \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{z}^{(k)}, \\ \mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k A \mathbf{z}^{(k)} \end{cases}$$

In questo caso stiamo introducendo un dinamismo ulteriore rispetto al metodo di Richardson. Mentre Richardson già prevedeva un parametro che permetteva di accelerare la convergenza, qui introduciamo un parametro che possiamo variare ad ogni iterazione.