RADICI DI EQUAZIONI NON LINEARI (Criteri di arresto per i metodi iterativi)

In questa lezione ci occuperemo di uno degli aspetti fondamentali dei metodi iterativi: determinare quando interrompere le iterazioni del metodo. Ovvero, individuare il minimo valore di n, l'indice di iterazione, a partire dal quale conviene arrestarsi e considerare l'ultima soluzione calcolata come un'approssimazione accettabile della soluzione cercata. In altre parole, definiremo quelli che vengono definiti criteri di arresto delle iterazioni.

Nelle lezioni precedenti, abbiamo studiato la determinazione degli zeri delle equazioni non lineari e dei sistemi di equazioni non lineari. In particolare, abbiamo analizzato diverse tecniche per definire metodi iterativi, che hanno lo scopo di costruire delle successioni di valori che, idealmente, tendano alla soluzione cercata.

Abbiamo visto il problema della determinazione delle radici di un'equazione non lineare $f(x)$

$$f(x) = 0,$$

dove $f(x)$ può essere un'equazione polinomiale o più in generale una funzione che fa intervenire delle non linearità rispetto alla variabile x. Abbiamo poi visto che questo concetto si può generalizzare al caso dei sistemi di equazioni non lineari

$$F(\boldsymbol{x}) = 0,$$

dove F in questo caso denoterà una funzione vettoriale e $\boldsymbol{x}$ un vettore. Infine, abbiamo visto il caso della ricerca dei punti fissi. Ovvero abbiamo cercato quei valori x della variabile indipendente in corrispondenza dei quali una certa funzione g assume valore esattamente uguale x

$$g(x) = x.$$

Per affrontare questi problemi matematici, abbiamo introdotto famiglie di metodi iterativi in grado di generare una successione di valori, siano essi numeri reali o vettori, tali che

$$\lim_{n\to\infty} x_n = x.$$

Un aspetto fondamentale è determinare il momento in cui interrompere il processo. In altre parole, individuare il minimo valore di n per cui possiamo considerare x_n sufficientemente accurato, ovvero abbastanza vicino alla soluzione x del problema dato. Per questo siamo interessati a valutare quando l'errore di approssimazione e_n, a un certo passo n , della soluzione approssimata x_n del metodo, rispetto alla soluzione esatta x, sia al di sotto di una soglia (o tolleranza) accettabile. Definiamo allora errore di approssimazione al passo n come

$$e_n = |x - x_n|,$$

dove x_n è la soluzione approssimata al passo n dal metodo e x la soluzione esatta. Vogliamo, quindi, essere in grado di controllare l'errore entro una certa soglia ϵ fissata (dove, ϵ può essere, ad esempio, $10^{-2}, 10^{-5}, 10^{-10}$, a seconda del livello di accuratezza desiderato per la soluzione del problema

$$e_n = |x - x_n| < \epsilon.$$

Osserviamo che nella maggior parte dei casi, non siamo in grado di conoscere a priori la soluzione esatta. Quindi abbiamo bisogno di introdurre dei criteri che ci permettano di controllare l'errore senza necessitare della conoscenza a priori della soluzione esatta.

Introduciamo di seguito due criteri per valutare il grado di convergenza di un metodo iterativo. Il primo è l'analisi del residuo, indicato con r_n. Definiamo residuo al passo n, per l'equazione $f(x) = 0$, il valore r_n t. c.

$$r_n = |f(x_n)|.$$

Nel caso della soluzione esatta x, vale la relazione $f(x) = 0$. Se x_n è abbastanza vicino a x, allora il residuo r_n sarà piccolo.
Analogamente, la definizione di residuo per problemi vettoriali (sistemi di equazioni non lineari) è data dal vettore

$$r^n = \|F(x_n)\|,$$

in un'opportuna norma vettoriale.
Nel caso delle iterazioni di punto fisso, ricordiamo che un valore x, detto punto fisso dell'applicazione g, soddisfa la relazione $g(x) = x$. In questo caso, il residuo è dato da

$$r_n = |g(x_n) - x_n|$$

Questo valore indica di quanto l'iterata corrente n si discosta dalla soluzione limite x, fornendo una misura dell'errore rispetto al soddisfacimento della condizione $g(x) = x$.
Vedremo di seguito, se l'analisi del residuo sarà significativa e in quali circostanze lo sarà. Ovvero in quali circostanze avere un residuo piccolo è garanzia di avere un errore piccolo; quindi, la differenza in valore assoluto tra la soluzione limite (o esatta) e la soluzione al passo n è piccola

$$e_n = |x - x_n| < \epsilon.$$

Come secondo criterio di controllo dell'errore introduciamo ora l'analisi incrementale, ovvero consideriamo come altro misuratore di errore, l'incremento δx

$$\delta x = |x_{n+1} - x_n|$$

ovvero la differenza in valore assoluto tra due iterate consecutive.
È evidente che, se x_n converge a x, ci si può aspettare che, a partire da un certo passo, la differenza δx diventi molto piccola. Osserveremo che monitorare questa differenza può essere utile in alcune situazioni, ma in altre non lo è. Infatti, possiamo costruire esempi in cui, nonostante un incremento molto piccolo, l'errore complessivo risulti essere in realtà molto grande.
Notiamo che l'incremento nel caso dei punti fissi è uguale al residuo

$$\delta x = |x_{n+1} - x_n| = |g(x_n) - x_n| = r_n,$$

dato che $x_{n+1} = g(x_n)$ per le iterazioni di punto fisso. Nel caso specifico dell'interazione di punto fisso, di fatto, i due criteri di arresto (analisi residuo e l'analisi dell'incremento) coincidono.

Di seguito ci accingiamo a fare un'analisi specifica dei singoli casi: metodi per la ricerca di radici di equazioni non lineari e iterazioni di punto fisso.
Concentrarci dapprima sul problema della ricerca di radici di equazioni non lineari

$$f(\alpha) = 0.$$

Cerchiamo la radice α (che questa volta abbiamo chiamato α e non più x perché, nella trattazione delle lezioni precedenti, abbiamo indicato con α la radice della funzione f) dell'equazione non lineare

$$f(x) = 0.$$

Costruiamo, utilizzando un generico metodo numerico, una successione $\{x_n\}$ tale che

$$\{x_n\} \Longrightarrow \lim_{n\to\infty} x_n = \alpha.$$

A tale scopo, abbiamo visto che possiamo utilizzare diversi metodi: la bisezione, le corde, le seccanti, il metodo di Newton. In tutti questi casi si introduce una successione di valori $\{x_n\}$ che converge sotto opportune condizioni alla radice α.
Il problema, dunque, è determinare quando fermare le iterazioni dei nostri metodi numerici. Per fare ciò, ci baseremo sui due criteri di arresto appena introdotti: il controllo del residuo, che in questo caso

corrisponde al valore assoluto di f(x_n) e il controllo dell'incremento, che rappresenta il valore assoluto di $x_{n+1} - x_n$, ovvero

$$r_n = |\, f(x_n)|,$$
$$\delta x = |x_{n+1} - x_n\,|.$$

Entrambe queste quantità sono direttamente calcolabili in funzione della dall'ultimo valore $n + 1$, o del penultimo n, che abbiamo precedentemente calcolato.
Ricordiamo che l'obiettivo è quello di avere un errore e_n piccolo

$$e_n = |\alpha - x_n|,$$

dove α è la soluzione esatta e x_n la soluzione al passo n.
Questo è il quadro per quanto riguarda i metodi per la ricerca degli zeri di equazioni non lineari. Ora vogliamo esaminare come questi criteri funzionano e in quali circostanze risultano efficaci.
Dunque, vediamo il controllo del residuo

$$r^n = |\, f(x_n)| < \epsilon.$$

L'idea è quella quindi di fermarsi non appena troviamo un valore di n tale che $|\, f(x_n)|$ sia minore di ϵ, dove ϵ rappresenta una tolleranza stabilita a priori.

Vediamo un esempio: sia f una generica funzione considerata e n l'ultima iterata calcolata. Il valore assoluto di $f(x_n)$ rappresenta la lunghezza del segmento che unisce la curva, in corrispondenza di x_n, all'asse delle x. D'altra parte, e_n è la distanza tra la soluzione esatta e la soluzione nell'iterata n. In questo caso, garantire un residuo piccolo è sufficiente per assicurarsi che l'errore sia piccolo. I due segmenti hanno, infatti, all'incirca la stessa lunghezza. Pertanto, se $|\, f(x_n)|$ è piccolo, anche e_n sarà piccolo. Tuttavia, esistono casi in cui, pur avendo un residuo piccolo, l'errore risulta comunque grande. Di conseguenza, si possono verificare due situazioni differenti: nella prima, il controllo del residuo sembra essere sufficiente per garantire anche un buon controllo sull'errore. Al contrario, nella seconda situazione, questo non è più vero. È infatti possibile costruire esempi di funzioni in cui, nonostante un residuo piccolo, l'errore al passo n-esimo è molto grande. Siamo indotti a voler comprendere il meccanismo matematico che regola questo tipo di comportamento. L'obiettivo è arrivare a una formula generale che ci permetta di trarre delle conclusioni non per un caso specifico, ma per tutti i casi.

Lo strumento che possiamo utilizzare è il teorema del valor medio. Iniziamo partendo da questa identità:

$$f(x_n) = f(x_n),$$

da cui sottraiamo uno zero (essendo α una radice e quindi $f(\alpha) = 0$), lasciando inalterata l'identità

$$f(x_n) = f(x_n) - f(\alpha).$$

Applichiamo adesso il teorema del valor medio

$$f(x_n) = f(x_n) - f(\alpha) = f'(\xi_n)(x_n - \alpha)$$

il quale mi dice che esiste un punto ξ_n compreso tra α e x_n tale per cui $f(x_n) - f(\alpha)$ è uguale a $f'(\xi_n)(\xi_n - \alpha)$. Allora, isolando $\alpha - x_n$, che è il nostro errore, otteniamo

$$\alpha - x_n = -\frac{f(x_n)}{f'(\xi_n)}.$$

Se adesso prendiamo i valori assoluti possiamo osservare che

$$e_n = |\alpha - x_n| = \left|-\frac{f(x_n)}{f'(\xi_n)}\right| \simeq \frac{1}{|f'(\alpha)|}|f(x_n)|.$$

In particolare, nel caso in cui f sia una funzione continua e derivabile con continuità, possiamo supporre che $f'(\xi_n) \simeq f'(\alpha)$, dato che x_n sta convergendo ad α ed anche ξ_n ,che è compreso tra x_n e α, sta convergendo ad α. Questo è il motivo per cui abbiamo messo un segno di circa uguale e non uguale.

Abbiamo quindi scritto una relazione fra il residuo $f(x_n)$ e l'errore e_n. Il parametro di possibili amplificazioni dell'errore è $\frac{1}{|f'(\alpha)|}$. Dunque, possiamo pensare che a un residuo piccolo corrisponda un errore piccolo, quando $|f'(\alpha)|$ sia grande. Se invece $|f'(\alpha)|$ fosse molto piccolo, ad un residuo piccolo potrebbe corrispondere un errore molto grande.

Le considerazioni fatte sui due esempi specifici, possiamo ritrovarle come casi particolari di questa formula generale. Esistono tre casi limite:

1. $f'(\alpha) \sim 1$
2. $f'(\alpha) \ll 1$ piccolo
3. $f'(\alpha) \gg 1$ grande

Caso 1: $f'(\alpha) \sim 1$. In questo caso il criterio sul residuo è un buon test. Infatti
$$e_n \simeq \frac{1}{|f'(\alpha)|}|f(x_n)| \sim |f(x_n)|,$$
ovvero, il valore dell'errore è circa uguale al valore del residuo. In questo caso, il test sul residuo è un buon test di arresto.
Caso 2: $f'(\alpha) \ll 1$. Caso in cui $f'(\alpha)$ sia molto piccolo, cioè molto minore di uno, circa uguale a zero. Allora in questo caso abbiamo un fattore di amplificazione importante. Possiamo avere un residuo $|f(x_n)|$ piccolo, ma un errore grande. Questo è il caso in cui $f'(\alpha)$ ha una derivata nel punto α che è molto vicina allo zero.
Caso 3: $f'(\alpha) \gg 1$. Nel caso in cui $f'(\alpha)$ sia molto grande, certamente a un residuo piccolo corrisponde un errore molto piccolo.
$$e_n \simeq \frac{1}{|f'(\alpha)|}|f(x_n)|.$$
Tuttavia, il test potrebbe risultare troppo restrittivo. Infatti, per una curva di forte pendenza, potrebbe accadere che, nonostante un residuo relativamente grande, l'errore sia effettivamente molto più piccolo della tolleranza richiesta. In questa situazione, il test ci costringerebbe a proseguire con le iterazioni, anche quando l'errore è già sufficientemente piccolo, ma il residuo continua a essere grande. Pertanto, questo test ci garantisce sicuramente che l'errore sia piccolo, ma potrebbe anche indurci a compiere iterazioni inutili, anche quando ci troviamo già in un intorno della soluzione che possiamo considerare molto soddisfacente.

Vogliamo ora esaminare come si comporta il test di arresto basato sull'incremento, ossia sulla differenza tra due iterazioni successive. Cambiamo quindi il tipo di test e analizziamo il controllo dell'incremento. In questo caso, ci fermiamo quando
$$\delta x = |x_{n+1} - x_n| < \epsilon.$$
Quindi, a una distanza $|x_{n+1} - x_n|$ può corrispondere un errore di valore all'incirca simile. Questo rappresenta un caso favorevole, in cui l'errore ha un valore che è approssimativamente equivalente al valore dell'incremento.

Una situazione non buona invece è quando ad un incremento piccolo può corrispondere un errore molto grande. Questo è un caso in cui il test non darebbe risultati soddisfacenti. Ci indurrebbe a fermarci in corrispondenza di un n per il quale abbiamo un incremento piccolo, ma per x_n l'errore sarebbe ancora molto grande.

Concludiamo andando ad analizzare come il criterio di arresto sull'incremento possa applicarsi al caso del metodo di Newton e delle iterazioni di punto fisso. In particolare, vedremo che per il metodo di Newton, il test sull'incremento è sempre un buon test. Quindi se le x_n sono ottenute per applicazione del metodo di Newton, allora il test sull'incremento funzionerà bene.

Ricordiamo che il metodo di Newton si scrive come

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

Allora, se utilizziamo il teorema del valor medio,

$$\alpha - x_n = -\frac{f(x_n)}{f'(\xi_n)},$$

nella relazione del metodo di Newton, troviamo

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

dove abbiamo approssimato $f'(x_n) \simeq f'(\xi_n)$, garantito dal fatto che f è una funzione continua e derivabile con continuità nel caso in cui x_n sia molto vicino a ξ_n. Quindi,

$$x_{n+1} - x_n \simeq \alpha - x_n,$$

che ci dice che c'è una corrispondenza quasi perfetta tra la l'incremento e l'errore effettivo.

Abbiamo quindi trovato che nel metodo di Newton l'errore al passo n è essenzialmente uguale al valore dell'incremento a quel passo. Dunque, nel caso del metodo di Newton, il criterio di arresto dell'incremento funzionerà bene perché ci garantirà appunto che ad un incremento piccolo corrisponderà un errore piccolo sulla soluzione.

Vediamo invece cosa succede nel caso delle iterazioni di punto fisso. Ricordiamo che le interazioni punto fisso sono definite dalla formula

$$x_{n+1} = g(x_n).$$

In questo caso, sfruttando questa relazione e applicando di nuovo il teorema del valor medio si ottiene

$$\alpha - x_n \simeq \frac{1}{1 - g'(\alpha)}(x_{n+1} - x_n).$$

Questo fattore $\frac{1}{1-g'(\alpha)}$ è il fattore di potenziale amplificazione dell'incremento. Ricordiamoci che le iterazioni di punto fisso convergono tanto più rapidamente quanto più $g'(\alpha)$ è piccolo ed in particolare

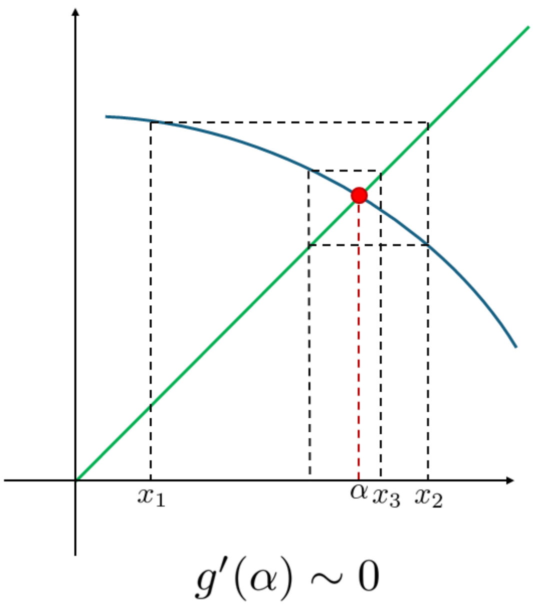

Figura 1: Iterazioni di punto fisso per un caso in cui $g'(\alpha) \sim 0$.

se $g'(\alpha) = 0$, le interazioni convergerebbero quadraticamente. Allora, se $g'(\alpha) \sim 0$ il test sull'incremento è un buon test perché ci dice che l'errore è circa uguale all'incremento, esattamente come succedeva nel caso del metodo di Newton. Tuttavia, se $g'(\alpha) \simeq 1$, quindi non piccolo ma dell'ordine dell'unità, il denominatore sarebbe circa uguale zero o comunque molto piccolo

$$\alpha - x_n \simeq \frac{1}{1 - g'(\alpha)}(x_{n+1} - x_n),$$

e dunque avremo un potenziale fattore di amplificazione dell'errore importante. Ovvero avremo che l'incremento è moltiplicato per questo fattore di amplificazione che potrebbe dare un errore grande. Quindi in questo caso non sarà conveniente utilizzare questo test per le iterazioni di punto fisso.

Vediamo adesso, su un esempio specifico, il caso $g'(\alpha) \sim 0$ (si veda Figura 1). La pendenza della curva in questo punto non è uguale a zero, ma non è molto grande. Alla prima iterazione di punto fisso stiamo già vedendo che l'errore $e_n = |\alpha - x_n|$ è più piccolo dell'errore sull'incremento $|x_{n+1} - x_n|$. Quindi, in particolare, se ci fermassimo per $|x_{n+1} - x_n|$ piccolo, avremo la certezza che anche

$$e_n = |\alpha - x_n|$$

sarebbe piccolo. E quindi questo verifica effettivamente la bontà del test sull'incremento.

Riassumendo, abbiamo quindi visto che, per quanto riguarda le equazioni non lineari, possiamo utilizzare due test di arresto: il test sul residuo al passo n e il test sull'incremento. Abbiamo analizzato le regole matematiche che permettono di mettere in relazione questi valori con l'errore sulle soluzioni, che è ciò che desideriamo effettivamente stimare come piccolo entro una certa soglia. Questa analisi l'abbiamo fatta per le iterazioni di punto fisso e per le iterazioni finalizzate alla ricerca degli zeri di equazioni non lineari. Potremmo ripetere esattamente lo stesso processo per il caso dei sistemi di equazioni non lineari, ovvero nel caso in cui F sia una funzione vettoriale. Le conclusioni sarebbero essenzialmente le stesse. Ci sarebbero delle differenze nel formalismo, perché dovremmo considerare, ad esempio, non i valori assoluti degli errori, ma le norme (di vettore) degli errori e dei residui. Inoltre, dovremmo applicare il teorema del valor medio non più al caso delle equazioni scalari, ma a quello dei sistemi di equazioni. Tuttavia, le conclusioni a cui siamo giunti rimarrebbero sostanzialmente invariate.