RADICI DI EQUAZIONI NON LINEARI (Iterazioni di punto fisso per la ricerca degli zeri e sistemi di equazioni non lineari)

In questa lezione analizzeremo come, in generale, sia possibile stabilire un'equivalenza tra la ricerca dei punti fissi di una funzione e la determinazione degli zeri di un'altra funzione. Nella seconda parte, mostreremo inoltre come estendere i metodi studiati in precedenza per la ricerca degli zeri al caso dei sistemi di equazioni non lineari.

Abbiamo già esaminato come definire iterazioni per ottenere la convergenza ai punti fissi di una funzione $g(x)$. Inoltre, nella lezione precedente, abbiamo esaminato un caso specifico in cui la ricerca dei punti fissi può essere ricondotta al calcolo degli zeri di una funzione.

Supponiamo di avere due funzioni f(x) e $g(x)$ legate dalla seguente relazione:

$$f(x) = x - g(x).$$

In questo caso, i punti fissi di $g(x)$ coincidono con le radici di $f(x)$, poiché:

$$f(x) = 0 \iff x = g(x).$$

Vediamo cosa accade del punto di vista geometrico. Abbiamo visto che trovare gli zeri di $f(x)$ equivale a determinare i punti di intersezione della funzione f con l'asse x. In questo contesto, f e g sono collegate nel seguente modo: i valori di x per cui $f(x) = 0$ corrispondono ai valori di x in cui il grafico di g interseca la bisettrice del primo e del terzo quadrante.

Dato che esiste una corrispondenza tra le radici di f e i punti fissi di g, possiamo reinterpretare i metodi già introdotti per il calcolo delle radici di f come metodi di iterazione del punto fisso, utili per approssimare i punti fissi di una funzione g associata a f.

In particolare, analizziamo il Metodo delle Corde e il Metodo di Newton, osservando come possano essere interpretati alla luce di questa relazione.

Ricordando che il metodo delle corde è definito da

$$x_{n+1} = x_n - \frac{b-a}{f(b)-f(a)} f(x_n) \qquad \forall n \geq 0,$$

e definendo

$$g(x_n) = x_n - \frac{b-a}{f(b)-f(a)} f(x_n),$$

allora possiamo caratterizzare il metodo delle corde come un metodo di punto fisso, ovvero

$$x_{n+1} = g(x_n) \qquad \forall n \geq 0.$$

Questo, dunque, è il Metodo delle Corde reinterpretato come un metodo di iterazione del punto fisso.

Analogamente possiamo definire il metodo Newton, come metodo di punto fisso. Il metodo di Newton è definito da

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \qquad \forall n \geq 0.$$

Se definiamo

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

allora

$$x_{n+1} = g(x_n) \qquad \forall n \geq 0.$$

Abbiamo rivisto il metodo delle corde e il metodo di Newton nell'ottica dei metodi di iterazione del punto fisso. Osserviamo ora che questa stessa interpretazione non è applicabile al metodo delle secanti. Infatti, il metodo delle secanti definisce x_{n+1} in funzione di due valori precedenti, x_n e x_{n-1}, ossia:

$$x_{n+1} = \bar{g}(x_n, x_{n-1}).$$

Di conseguenza, il metodo delle secanti non può essere espresso nella forma $x_{n+1} = g(x_n)$, tipica dei metodi del punto fisso. Per questo motivo, il metodo delle secanti non rientra tra i metodi di iterazione del punto fisso.

Abbiamo visto che il metodo Newton può essere ridefinito come metodo di punto fisso, osservando che esso è definito da

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \qquad \forall n \geq 0,$$

e definendo

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

da cui

$$x_{n+1} = g(x_n) \qquad \forall n \geq 0.$$

Andiamo adesso a verificare come per la funzione g associata al metodo di Newton si abbia che

$$g'(\alpha) = 0.$$

Infatti, andando a derivare

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

si ricava che

$$g'(x) = 1 - \frac{[f'(x)]^2 - f(x)f''(x)}{[f'(x)]^2},$$

da cui

$$g'(\alpha) = 1 - \frac{[f'(\alpha)]^2 - f(\alpha)f''(\alpha)}{[f'(\alpha)]^2},$$

Dato che $f(\alpha) = 0$, poiché α è uno zero della funzione f , la relazione precedente si riduce a

$$g'(\alpha) = 1 - \frac{[f'(\alpha)]^2}{[f'(\alpha)]^2} = 1 - 1 = 0.$$

Quindi abbiamo verificato che $g'(\alpha) = 0$ nel caso dal metodo di Newton. Questo ci permette di stabilire un anello di congiunzione tra proprietà teoriche di convergenza del metodo di Newton e quelle teoriche di convergenza dei punti fissi. Vedremo in dettaglio queste proprietà nella prossima lezione. Anticipiamo in questa lezione che il metodo di Newton converge "velocemente" alla radice della funzione: il metodo di Newton ha una convergenza quadratica (ovvero, il metodo di Newton è un metodo del secondo ordine) come vedremo. Mentre le generiche iterazioni di punto fisso convergono "più lentamente" rispetto al metodo di Newton: osserveremo che, in generale, i metodi di punto fisso hanno convergenza lineare (ovvero, i metodi di punto fisso sono metodi di primo ordine). Tuttavia, si riesce a dimostrare che se

$$g'(\alpha) = 0,$$

le interazioni di punto fisso convergono quadraticamente alla soluzione $g(\alpha) = \alpha$.

Quindi da un lato Newton ha una convergenza quadratica, dall'altro abbiamo detto che le interazioni di punto fisso convergono quadraticamente ogni volta in cui $g'(\alpha) = 0$. Quindi Newton è visto come metodo per la ricerca dei punti fissi e deve essere un metodo secondo ordine, perché $g'(\alpha) = 0$.

In questa seconda parte della lezione, desideriamo analizzare come i metodi introdotti per la ricerca degli zeri di funzioni non lineari possano essere generalizzati a problemi più complessi, in particolare al caso di problemi vettoriali.

Supponiamo che $f(x)$ non sia più una funzione scalare reale di variabile reale, ma una funzione vettoriale. Siamo interessati quindi a studiare come affrontare il problema della risoluzione di sistemi di equazioni non lineari e come applicare i metodi già visti per trovare le loro soluzioni.
Consideriamo F come una funzione vettoriale di variabile vettoriale. L'obiettivo è trovare un vettore $\boldsymbol{\alpha} \in I \subset \mathbb{R}^n$ tale che $F(\boldsymbol{\alpha}) = 0$

$$F: I \subset \mathbb{R}^n \to \mathbb{R}^m$$

$$\text{si cerca } \alpha \in I \subset \mathbb{R}^n \text{ t.c. } F(\boldsymbol{\alpha}) = 0$$

In altre parole, F è definita in un sottoinsieme dello spazio $\mathbb{R}^n$ e ha valori in $\mathbb{R}^m$, con $n > 1$ e $m > 1$ arbitrari. Il problema consiste quindi nel determinare il vettore $\boldsymbol{\alpha}$, composto da n componenti, che soddisfa l'equazione

$$F(\boldsymbol{\alpha}) = 0$$

Facciamo un esempio con $n = m = 2$. In questo caso, la funzione vettoriale F è definita come

$$F = (f_1, f_2),$$

con

$$F: I \subset \mathbb{R}^2 \to \mathbb{R}^2$$

Quindi, F è una funzione vettoriale di variabile vettoriale con due componenti scalari f_1 e f_2. Questo significa che F è definita nel piano bidimensionale e assume valori anch'essa nel piano bidimensionale.
Osserviamo che siamo alla ricerca dei valori α_1 e α_2 tali per cui

$$F(\alpha_1, \alpha_2) = (0,0).$$

Dal punto di vista geometrico, questo significa determinare i punti (x_1, x_2) nel piano delle variabili indipendenti per cui entrambe le componenti della funzione vettoriale, $f_1(x_1, x_2)$ e $f_2(x_1, x_2)$, si annullano simultaneamente.
In altre parole, stiamo cercando l'intersezione dei luoghi geometrici definiti dalle equazioni

$$f_1(x_1, x_2) = 0 \qquad f_2(x_1, x_2) = 0,$$

nel piano $x_1|x_2$, ovvero di fatto i punti che rappresentano le soluzioni del sistema non lineare.
La funzione f_1, essendo una funzione di due variabili a valori reali, definisce una superficie nello spazio tridimensionale, e lo stesso vale per f_2. Per trovare la soluzione del sistema $F(\boldsymbol{\alpha}) = 0$, consideriamo l'intersezione di queste due superfici nel piano $x_1|x_2$. Questa intersezione genera due curve corrispondenti agli insiemi di punti in cui rispettivamente $f_1(x_1, x_2) = 0$ e $f_2(x_1, x_2) = 0$. La soluzione del problema $F(\boldsymbol{\alpha}) = 0$ è quindi data dai punti di intersezione di queste due curve, ossia i valori di (x_1, x_2) per cui sia f_1 che f_2 si annullano contemporaneamente. In termini vettoriali, ciò equivale a dire che

$$F(x_1, x_2) = 0.$$

Più precisamente, le soluzioni saranno i punti che rappresentano le radici del sistema.

L'esempio considerato riguarda il caso bidimensionale, ma se consideriamo una funzione vettoriale definita su uno spazio di dimensione arbitraria, l'interpretazione geometrica diventa più complessa, poiché si tratta di intersezioni di iper-superfici in spazi di dimensione superiore.

Vogliamo ora mostrare come riprendere completamente il formalismo utilizzato nel caso delle funzioni reali di variabile reale e come estenderlo al caso vettoriale, al fine di generare una successione di iterazioni che, si spera, converga alla soluzione del problema

$$\{\boldsymbol{x}_k\} \implies \lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{\alpha}.$$

Vediamo ora come definire, ad esempio, il metodo di Newton per sistemi di equazioni non lineari. Ricordiamo che, nel caso delle equazioni scalari, il metodo di Newton è definito dalla relazione:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \qquad \forall k \geq 0,$$

che possiamo riscrivere come

$$f'(x_k)(x_{k+1} - x_k) = -f(x_k) \qquad \forall k \geq 0.$$

Questa non è altro che una riformulazione del metodo di Newton. Ora possiamo applicare lo stesso approccio al caso di problemi vettoriali, cioè quando F rappresenta un'equazione vettoriale. Riprendiamo quindi il formalismo e riscriviamo la formula precedente al caso vettoriale:

$$F'(\boldsymbol{x}_k)(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k) = -F(\boldsymbol{x}_k) \qquad \forall k \geq 0,$$

dove $F(\boldsymbol{x}_k)$ è un vettore e $\boldsymbol{x}_{k+1} - \boldsymbol{x}_k$ è anche esso un vettore (perché la differenza tra due vettori). Di conseguenza, la derivata $F'(\boldsymbol{x}_k)$ deve essere una matrice. In particolare, si tratta della matrice Jacobiana di F. La matrice Jacobiana è la matrice che ha come componenti le derivate parziali delle componenti di F, cioè delle componenti $f_1, f_2, \dots, f_m$, rispetto alle variabili $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$. Pertanto, trovare $F'(\boldsymbol{x}_k)$, significa calcolare questa matrice e valutarla nel punto corrispondente al vettore $(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n)$, che si suppone noto.

Per ogni k, abbiamo un sistema lineare da risolvere, poiché $F'(\boldsymbol{x}_k)$ è una matrice che dipende da $\boldsymbol{x}_k$. Ad ogni iterazione, quindi, ci troviamo a risolvere un sistema con la matrice jacobiana $F'(\boldsymbol{x}_n)$, dove il vettore $\boldsymbol{x}_{n+1} - \boldsymbol{x}_n$ (detto anche vettore incremento) è l'incognita e il vettore a destra dell'uguaglianza (detto vettore residuo) è il termine noto. Pertanto, ad ogni passo dell'iterazione, dobbiamo risolvere un sistema lineare.

Concludiamo questa lezione rimarcando che i metodi introdotti per la ricerca degli zeri di funzioni non lineari possano essere generalizzati a problemi più complessi, in particolare al caso di problemi vettoriali. Abbiamo visto come estendere il metodo di Newton per approssimare la soluzione di sistemi di equazioni non lineari. In particolare, per una funzione vettoriale F di variabile vettoriale x, trovare un vettore α tale che $F(\alpha) = 0$, riconduce alla risoluzione
di un sistema lineare, per ogni iterazione k, del tipo

$$F'(x_k)(x_{k+1} - x_k) = -F(x_k) \qquad \forall k \geq 0.$$

La risoluzione del problema $F(\alpha) = 0$, non è quindi affatto banale, perché ad ogni passo dobbiamo trattare una nuova matrice, $F'(x_k)$, che dipende da k. Calcolare questa matrice può essere molto dispendioso, soprattutto se il numero di iterazioni è elevato. Per questo motivo, nella pratica, talvolta si preferisce calcolare $F'(x_k)$ (detta la matrice Jacobiana) una sola volta, ad esempio in corrispondenza del vettore iniziale x_0. In questo caso, definendo A come

$$A = F'(x_0),$$

questa matrice A rimane costante durante le iterazioni. In pratica, ad ogni passo della procedura, usiamo sempre la stessa matrice A. Questa strategia è equivalente al metodo delle corde, dove la pendenza q_n della retta viene fissata. Qui, la pendenza è espressa dalla matrice Jacobiana, ma viene trattata come costante, similmente alla pendenza fissa nel Metodo delle Corde. Possiamo fare qualcosa di meglio, aggiornando la matrice A, ogni p iterazioni. In questo caso, calcoliamo A inizialmente al passo 0, la manteniamo costante per p iterazioni, e poi, a partire dall'iterazione $p + 1$, ricalcoliamo la matrice jacobiana A nel nuovo vettore x_p. Questo approccio è noto come metodo Quasi-Newton, poiché cerca di simulare il Metodo di Newton. Tuttavia, invece di utilizzare la Jacobiana che rappresenta la pendenza effettiva della nostra iper-superficie (o piano tangente), il metodo Quasi-Newton sostituisce la Jacobiana con una matrice fissa, ottenuta attraverso un aggiornamento periodico.