



## CALCOLO DEGLI AUTOVALORI E FONDAMENTI DELLA MATEMATICA NUMERICA (Condizionamento dei problemi numerici)

In questa In questa lezione affronteremo il problema della determinazione del numero di condizionamento relativo a un problema numerico. Verranno presentati alcuni esempi esplicativi. Analizzeremo, inoltre, l'incidenza del condizionamento numerico nei processi che implicano l'utilizzo di tali operazioni discrete nel calcolatore.

Abbiamo definito il numero di condizionamento di un problema numerico  $K_n$  come

$$K_n = \sup_{\delta d_n} \left( \frac{\|\delta u_n\|}{\|u_n\|} / \frac{\|\delta d_n\|}{\|d_n\|} \right),$$

ovvero, l'estremo superiore del rapporto tra la variazione nelle soluzioni  $\frac{\|\delta u_n\|}{\|u_n\|}$  in rapporto alla variazione nei dati  $\frac{\|\delta d_n\|}{\|d_n\|}$ . Tali rapporti sono intesi in senso relativo, ovvero calcolando le variazioni relative delle soluzioni divise per le variazioni relative dei dati, all'interno di una norma opportuna. Abbiamo inoltre affermato che, fissato un valore di  $n$ , si considera il numero di condizionamento  $K_n$  e si prende l'estremo superiore al variare di tutti gli  $n$  maggiori di  $k$ , ottenendo così una quantità dipendente da  $k$ . Considerando quindi il limite di tale quantità al tendere di  $k$  all'infinito, se tale limite esiste, esso viene indicato con  $K^{num}$  e denominato condizionamento numerico asintotico

$$K^{num} = \lim_{k \rightarrow \infty} \sup_{n \geq k} K_n.$$

A questo punto, ci proponiamo di fornire delle espressioni specifiche per il calcolo del numero di condizionamento numerico  $K_n$ . Ricordiamo che, se

$$F_n(u_n, d_n) = 0,$$

indica il problema numerico, talvolta è possibile rappresentare la soluzione  $u_n$  come una funzione  $f_n$  esplicita di  $d_n$

$$u_n = f_n(d_n).$$

Qualora ciò sia possibile, e nel caso in cui  $f_n$  sia una funzione differenziabile, ripetendo esattamente il procedimento adottato per il problema matematico al fine di ottenere  $K$ , ossia il numero di condizionamento del problema matematico, si perviene a un'espressione del tutto analoga

$$K_n \simeq \|f'_n(d_n)\| \frac{\|d_n\|}{\|f_n(d_n)\|},$$

dove  $K_n$  risulta approssimativamente uguale al prodotto tra la norma della derivata di  $f_n$ , calcolata in  $d_n$ , e il rapporto tra la norma di  $d_n$  e la norma di  $f_n(d_n)$ . Si ricorda che il simbolo  $\simeq$  indica un'approssimazione, che implica la trascurabilità di termini infinitesimi di ordine superiore rispetto a  $d_n$ .

Analizziamo ora come applicare la formula precedentemente introdotta in alcuni casi specifici, al fine di determinare quale forma assuma la funzione  $f_n$  e, di conseguenza, come caratterizzare quantitativamente il valore di  $K_n$ .

Il primo esempio, di natura estremamente semplice e con finalità puramente illustrative rispetto alla formula introdotta, riguarda la somma di due numeri. Supponiamo di dover calcolare la somma tra due numeri reali  $a$  e  $b$



$$u = a + b = f(d),$$

il cui risultato è  $u$ , che possiamo chiamare  $f(d)$ , dove  $d$  è l'insieme dei dati  $a$  e  $b$

$$d = [a, b]^T.$$

Quindi  $a$  e  $b$  sono noti e  $u$  è il valore da calcolare. In questo caso  $f'(d)$ , essendo  $f(d)$  una funzione di variabile vettoriale  $d$  e di componenti a valori reali  $a$  e  $b$ , sarà il vettore di jacobiano

$$f'(d) = \left[ \frac{\partial f}{\partial a}, \frac{\partial f}{\partial b} \right]^T = [1, 1]^T,$$

ovvero, il vettore costituito dalle derivate parziali della funzione  $f$  rispetto ad  $a$  e  $b$ . Ne consegue che  $f'(d)$  è un vettore. La derivata di  $f$  rispetto ad  $a$  è pari a 1, e analogamente la derivata rispetto a  $b$  è anch'essa pari a 1. Inoltre, ricordiamo che, per un vettore generico  $v$ , con componenti  $v_1$  e  $v_2$ , una possibile norma è data dalla somma dei valori assoluti delle componenti, ovvero la cosiddetta norma 1

$$\|v\|_1 = |v_1| + |v_2|.$$

Possiamo dunque utilizzare la norma 1 come norma per calcolare il numero di condizionamento in questo caso di esempio

$$K \simeq \|f'(d)\| \frac{\|d\|}{\|f(d)\|} = 2 \frac{(|a| + |b|)}{|a + b|},$$

dove  $\|f'(d)\|$  è la norma 1 del vettore  $f'(d) = [1, 1]^T$ , ovvero la somma delle componenti  $1 + 1$ , la norma  $\|d\|$  è il valore assoluto di  $a$  più valore assoluto  $b$ , e la norma  $\|f(d)\|$ , essendo  $f(d)$  un numero reale, è uguale ad  $a$  più  $b$  in valore assoluto. Dunque, il numero di condizionamento in questo caso di esempio è

$$K \simeq 2 \frac{(|a| + |b|)}{|a + b|}.$$

Diremo che l'operazione di somma e di addizione è ben condizionata se  $a$  e  $b$  hanno lo stesso segno. Infatti, se  $a$  e  $b$  hanno lo stesso segno, allora  $a + b$  è un numero certamente diverso da zero e  $K$  è dell'ordine dell'unità (in particolare in questo caso è dell'ordine di 2)

$$K \simeq 2,$$

perché, avendo  $a$  e  $b$  lo stesso segno, il valore assoluto di  $a + b$  è uguale al valore assoluto di  $a$  più il valore assoluto di  $b$ . Il problema sarà mal condizionato quando

$$b \simeq -a,$$

ovvero, quando  $a$  e  $b$  hanno segno opposto ed inoltre  $b$  è circa uguale in modulo ad  $a$ . In questo caso, la somma  $a + b$  assume un valore molto prossimo allo zero. Pertanto, nell'applicazione della formula

$$K \simeq 2 \frac{(|a| + |b|)}{|a + b|},$$

il numeratore viene diviso per una quantità che è approssimativamente nulla. Di conseguenza, il valore di  $K$  risulta molto elevato e, in termini formali, tende all'infinito nel caso in cui  $b = -a$ , ovvero

$$K \simeq \infty.$$

Naturalmente, tale espressione ha carattere puramente formale e non possiede un significato effettivo in senso numerico rigoroso. Essa indica semplicemente che il rapporto considerato non è più definito nel caso in cui  $b$  sia esattamente uguale  $-a$ .

Si tratta di un processo critico, in quanto conduce a un numero di condizionamento particolarmente elevato. Ciò avviene quando si desidera sommare due numeri la cui relazione è tale che  $a + b \approx 0$ , ovvero quando  $b$  è approssimativamente pari all'opposto di  $a$ . Questa operazione, come vedremo a breve, può condurre, nell'ambito del calcolo numerico, al cosiddetto fenomeno di cancellazione di cifre significative. Infatti, se il denominatore  $a + b$  viene calcolato con un numero limitato di cifre



significative, il risultato potrebbe essere molto vicino allo zero, generando così un numero di condizionamento estremamente elevato.

Procediamo ora con alcuni esempi. Il primo riguarda la somma tra due numeri reali  $a$  e  $b$ , dove

$$a = \pi = 3.14159265 \dots, \quad b = -22/7 = -3.142857.$$

Qualora si utilizzassero soltanto cinque cifre significative per rappresentare questi due numeri - e vedremo a breve che un calcolatore opera effettivamente con un numero finito di cifre significative - si avrebbe che

$$a_M = 3.1416, \quad b_M = -3.1429,$$

dove  $a_M$  e  $b_M$  indicano, rispettivamente, la rappresentazione su calcolatore (detta rappresentazione di macchina) dei numeri  $a$  e  $b$ . Vedremo nel dettaglio la rappresentazione di macchina nella prossima lezione. Per ora ci basti osservare che per  $a_M = 3.1416$  viene approssimata la quarta cifra decimale da 5 a 6, poiché la cifra successiva è 9, valore maggiore o uguale a 5. Analogamente, si ha che  $b_M = -3.1429$  poiché la quarta cifra decimale è stata approssimata da 8 a 9, dato che la cifra successiva è 5. Abbiamo quindi a disposizione i numeri esatti  $a$  e  $b$ , e le rispettive approssimazioni in macchina,  $a_M$  e  $b_M$ , nell'ipotesi di un calcolatore che operi con cinque cifre significative.

Valutiamo gli errori introdotti sui dati. Invece di utilizzare i valori esatti  $a$  e  $b$ , impieghiamo le loro rappresentazioni in macchina  $a_M$  e  $b_M$ . Calcolando la differenza relativa, ovvero l'errore relativo, si riscontrano errori sui dati che risultano relativamente contenuti

$$\frac{|a - a_M|}{|a|} = 2.34 \cdot 10^{-6}, \quad \frac{|b - b_M|}{|b|} = 1.36 \cdot 10^{-5}.$$

Se ora calcoliamo l'errore relativo sul risultato, confrontando la differenza tra il valore esatto dell'operazione  $u = a + b$  e quello ottenuto mediante il calcolatore con cinque cifre significative,  $u_M = a_M + b_M$ , si ottiene che

$$\frac{|(a + b) - (a_M + b_M)|}{|a + b|} = 0.28 \cdot 10^{-1},$$

ovvero, l'errore riscontrato risulta essere dell'ordine di  $10^{-1}$ . Ciò significa che si è commesso un errore di tale entità, nonostante gli errori sui dati fossero di gran lunga più piccoli. Si è dunque verificata una significativa propagazione degli errori, fenomeno comunemente denominato, nella letteratura anglosassone, loss of significance error, ovvero perdita di cifre significative.

Quando si intende rappresentare numeri reali mediante valori troncati o arrotondati, è possibile che errori di entità ridotta, introdotti durante tale approssimazione, si traducano in errori di gran lunga maggiori nelle operazioni che coinvolgono tali valori approssimati. Ciò si verifica in particolare nelle operazioni di addizione, nel caso in cui i due numeri siano approssimativamente opposti tra loro.

Esaminiamo ora un secondo esempio di calcolo del numero di condizionamento, questa volta relativo ai sistemi lineari. Ricordiamo che un sistema lineare rappresenta una relazione del tipo

$$A\mathbf{u} = \mathbf{b},$$

dove  $A$ , la matrice dei coefficienti, è una matrice  $n \times n$ ;  $\mathbf{b}$ , il vettore dei termini noti, è un vettore con  $n$  componenti; e  $\mathbf{u}$ , il vettore incognito, rappresenta la soluzione del corrispondente sistema lineare. La funzione  $F(\mathbf{u}, d)$  in questo caso è

$$F(\mathbf{u}, d) = A\mathbf{u} - \mathbf{b},$$

e il problema matematico è

$$F(\mathbf{u}, d) = \mathbf{0}.$$

Se la matrice  $A$  è invertibile, la soluzione del sistema lineare, è

$$\mathbf{u} = A^{-1}\mathbf{b},$$



di conseguenza

$$\mathbf{u} = A^{-1}\mathbf{b} = f(d),$$

dove  $f$  è la relazione funzionale che fa corrispondere all'insieme dei dati  $d$

$$d = \{A, \mathbf{b}\},$$

la soluzione  $\mathbf{u}$  del sistema lineare  $A\mathbf{u} = \mathbf{b}$ .

Supponiamo ora di perturbare unicamente il termine noto  $\mathbf{b}$ . Supponiamo, quindi, di restringere i dati al solo insieme  $\mathbf{b}$

$$d = \{\mathbf{b}\},$$

ovvero, supponiamo di introdurre un errore nella rappresentazione del solo termine noto  $\mathbf{b}$ , senza commettere alcuna imprecisione nella rappresentazione della matrice  $A$ . In tal caso, il numero di condizionamento, applicando la formula generale

$$K \simeq \|f'(d)\| \frac{\|d\|}{\|f(d)\|},$$

risulta

$$K \simeq \frac{\|A^{-1}\| \|\mathbf{b}\|}{\|\mathbf{u}\|} = \frac{\|A^{-1}\| \|A\mathbf{u}\|}{\|\mathbf{u}\|} \leq \frac{\|A^{-1}\| \|A\| \|\mathbf{u}\|}{\|\mathbf{u}\|} = \|A^{-1}\| \|A\|,$$

dove  $f'(d)$  rappresenta la derivata prima  $\mathbf{u} = A^{-1}\mathbf{b} = f(d)$  rispetto al dato  $d = \mathbf{b}$ , da cui segue che  $f'(d) = A^{-1}$ . Il dato  $d$  corrisponde dunque al vettore  $\mathbf{b}$ , mentre la funzione  $f(d)$  rappresenta la soluzione  $\mathbf{u}$ . Inoltre, nel secondo passaggio si è utilizzata la relazione  $\mathbf{b} = A\mathbf{u}$ . Il segno di disuguaglianza deriva dalla relazione

$$\|A\mathbf{u}\| \leq \|A\| \|\mathbf{u}\|,$$

la quale afferma che la norma del prodotto di due operatori (o matrici) è minore o uguale al prodotto delle norme dei singoli fattori. Infine, semplificando la norma  $\|\mathbf{u}\|$  al numeratore e al denominatore, si ottiene

$$K(A) \simeq \|A^{-1}\| \|A\|.$$

Possiamo riassumere quanto emerso affermando che, in questo caso, il numero di condizionamento è dato dal prodotto tra la norma di  $A^{-1}$  e la norma di  $A$ . Questo rappresenta, quindi, il numero di condizionamento associato a un sistema lineare, ovvero il fattore di amplificazione potenziale degli errori presenti nei dati relativi al solo termine noto sulla soluzione del sistema.

Il valore  $K(A)$  ottenuto possiede un significato intrinseco e, più precisamente, prende il nome di numero di condizionamento della matrice  $A$ . Naturalmente, poiché non è stata specificata una particolare norma, essa può essere scelta arbitrariamente. Di conseguenza, la definizione del numero di condizionamento dipende dalla norma di matrice che si decide di adottare.

Osserviamo infine che il numero di condizionamento  $K(A)$  di una matrice  $A$  è sempre maggiore o uguale a uno

$$K(A) \geq 1,$$

poiché il numero di condizionamento della matrice identità  $I$  è pari a uno

$$K(I) = 1.$$

Infatti, la quantità  $\|A^{-1}\| \|A\|$  è per definizione maggiore o uguale alla norma del prodotto  $\|A^{-1}A\|$

$$\|A^{-1}A\| \geq \|A^{-1}\| \|A\|,$$

e poiché  $A^{-1}A = I$  e  $\|I\| = 1$ , si ottiene

$$1 = \|I\| = \|A^{-1}A\| \leq \|A^{-1}\| \|A\| = K(A).$$

La matrice  $A$  si dirà mal condizionata se il numero di condizionamento  $K(A)$  risulta molto maggiore di 1. Al contrario, se  $K(A)$  è dell'ordine dell'unità, si dirà che la matrice  $A$  è ben condizionata.



$K(A) \gg 1$ : matrice mal condizionata,  
 $K(A) \approx 1$ : matrice ben condizionata.

Possiamo quindi concludere che, se una matrice  $A$  è ben condizionata, la risoluzione del sistema lineare ad essa associato rappresenta un problema ben condizionato. Viceversa, se la matrice  $A$  è mal condizionata, la risoluzione del sistema lineare corrispondente risulta essere un problema mal condizionato. In altre parole, a priori, piccole variazioni nei dati del problema possono comportare variazioni significative nella soluzione stessa.

Consideriamo ora un terzo esempio, relativo al problema della ricerca degli zeri di un'equazione non lineare, che indichiamo con

$$\phi(x) = d,$$

dove  $d$  è un dato e  $\phi$  è una funzione non lineare della variabile  $x$ . In particolare,  $d$  rappresenta la retta parallela all'asse delle ascisse, posta a distanza  $d$  dall'asse  $x$ . La relazione  $\phi(x) = d$  indica la ricerca dell'ascissa  $x$  del punto di intersezione tra la curva  $\phi$  e la retta parallela all'asse delle ascisse.

Se  $\phi$  è una funzione invertibile, possiamo considerare  $x$  come funzione di  $d$  mediante la funzione inversa  $\phi^{-1}$

$$x = \phi^{-1}(d).$$

Si tratta quindi di un problema che può essere rappresentato mediante le notazioni convenzionali

$$F(u, d) = 0 \Leftrightarrow u = f(d),$$

ovvero, possiamo esprimere

$$x = \phi^{-1}(d) \Leftrightarrow u = f(d),$$

dove  $f = \phi^{-1}$  è l'inversa della funzione  $\phi$  e  $u = x$  è la soluzione incognita. Ricordiamo che

$$(\phi^{-1})'(d) = [\phi'(x)]^{-1},$$

ossia la derivata di  $\phi^{-1}$  in un punto  $d$  coincide con l'inversa della derivata di  $\phi$  nel punto  $x$  corrispondente a  $d$ . Tale relazione è fondamentale in quanto ci consente di calcolare il numero di condizionamento relativo al problema del calcolo degli zeri di un'equazione non lineare. Questo numero di condizionamento può essere espresso mediante la formula consueta

$$K \approx \|f'(d)\| \frac{\|d\|}{\|f(d)\|},$$

in cui  $f'(d) = [\phi'(x)]^{-1}$  (come derivato in precedenza),  $d = \phi(x)$  ed infine  $f(d)$  è  $x = f(d)$

$$K \approx \left| (\phi'(x))^{-1} \right| \frac{\|\phi(x)\|}{\|x\|}.$$

Il problema della ricerca degli zeri di una funzione non lineare si definisce mal condizionato quando  $\phi'(x)$  assume valori prossimi a zero, poiché in tal caso l'inversa  $(\phi'(x))^{-1}$  risulta molto grande, determinando un numero di condizionamento  $K$  elevato. Al contrario, il problema si considera ben condizionato quando  $\phi'(x)$  è grande. Infine, il problema poi è mal posto se  $x$  ha una radice multipla dell'equazione  $\phi(x) = d$ , ovvero, se  $\phi'(x) = 0$ . Infatti, una radice multipla implica che la derivata prima di  $\phi$  nella soluzione  $x$  sia nulla.

Riassumendo

$\phi'(x)$  piccolo: problema mal condizionato,  
 $\phi'(x)$  grande: problema ben condizionato,  
 $\phi'(x) = 0$ : problema mal posto.

La Figura 1 illustra graficamente questo comportamento, considerando una funzione non lineare  $\phi(x)$ . Si osserva che il problema risulta mal posto quando è presente una radice multipla, ovvero quando la derivata prima  $\phi'$  nel punto  $x$  è nulla.

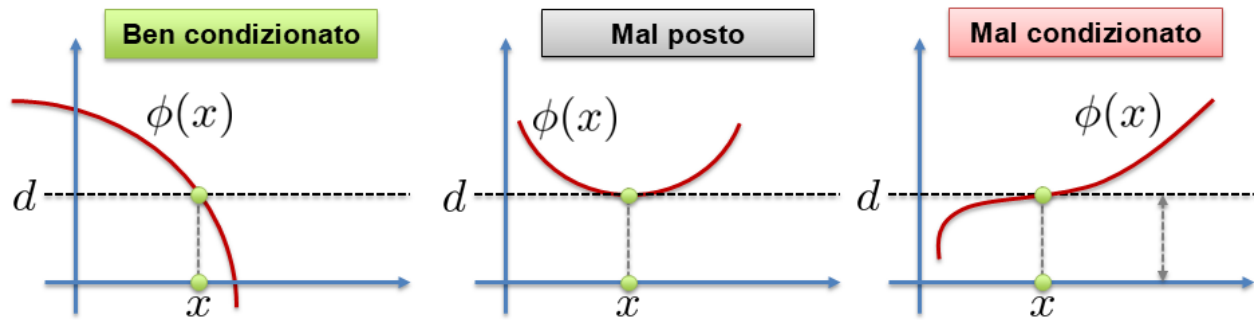


Figura 1: Rappresentazione grafica del condizionamento del problema della ricerca degli zeri per una funzione non lineare  $\phi(x)$ .

Il problema si definisce mal condizionato quando  $\phi'$ , pur non essendo nulla nel punto  $x$ , assume valori molto piccoli. Al contrario, il problema è ben condizionato quando  $\phi'$  è significativamente diverso da zero nel punto  $x$ .

Consideriamo un caso particolare di questo problema, relativo alle equazioni di secondo grado. In questo caso, l'equazione  $\phi(x) = d$  si riduce semplicemente a un'equazione di secondo grado. L'obiettivo è verificare, nel caso delle equazioni di secondo grado, il comportamento generale precedentemente analizzato e confermare le conclusioni a cui si è giunti nell'analisi del problema generale  $\phi(x) = d$ .

Consideriamo, per esempio, l'equazione

$$x^2 - 2dx + 1 = 0,$$

le cui radici sono

$$x_{\pm} = d + \sqrt{d^2 - 1}, \quad x_{-} = d - \sqrt{d^2 - 1},$$

In particolare, la radice è doppia se  $d = 1$  oppure  $d = -1$ . In questi casi

$$d = 1: x^2 - 2x + 1 = 0 \Rightarrow x = 1,$$

$$d = -1: x^2 + 2x + 1 = 0 \Rightarrow x = -1.$$

Infatti, nei casi in cui  $d = \pm 1$ , la radice è doppia, poiché in entrambe le situazioni si ottengono quadrati perfetti. Nel primo caso, la radice è  $x = 1$ , mentre nel secondo caso è  $x = -1$ . La situazione  $d = \pm 1$  rappresenta il caso limite, ovvero il caso in cui possiamo aspettarci problemi. Consideriamo invece il caso in cui non siano previsti problemi, ossia quando il valore assoluto di  $d$  sia molto maggiore di uno, ovvero

$$|d| \gg 1.$$

In questo caso  $|d| \gg 1$ , le due radici  $x_{\pm}$  sono ben separate. Abbiamo infatti

$$x_{+} = d + \sqrt{d^2 - 1},$$

$$x_{-} = d - \sqrt{d^2 - 1},$$

e poiché  $d$  è molto grande, nel calcolo di  $x_{+}$  si somma a  $d$  un valore approssimativamente dello stesso ordine di grandezza di  $d$ , quindi

$$x_{+} = d + \sqrt{d^2 - 1} \simeq d + \sqrt{d^2} = d + d = 2d,$$

mentre il calcolo di  $x_{-}$  risulta circa uguale a zero, perché stiamo sottraendo a  $d$  qualcosa che approssimativamente è dello stesso ordine di grandezza di  $d$

$$x_{-} = d - \sqrt{d^2 - 1} \simeq d - \sqrt{d^2} = d - d = 0.$$

Abbiamo dunque un problema che, dal punto di vista matematico, risulta ben condizionato. Tuttavia, il metodo numerico impiegato per la sua risoluzione può risultare instabile rispetto agli errori di





arrotondamento. In particolare, nel calcolo della radice  $x_-$  utilizzando l'aritmetica finita - ossia approssimando con un numero limitato di cifre significative - si sottraggono due numeri di valore molto simile tra loro. Tale situazione rappresenta il classico esempio di somma tra due numeri quasi opposti, che, come visto anche in precedenza, può causare il fenomeno noto come cancellazione di cifre significative, il quale può generare instabilità nell'aritmetica finita.

Questo rappresenta, quindi, un esempio di problema matematicamente ben condizionato che può tuttavia dar luogo a problematiche numeriche potenzialmente instabili. Tuttavia, proprio poiché il problema è ben condizionato dal punto di vista matematico, è possibile sviluppare metodi numerici per il calcolo di tali soluzioni che siano essi stessi ben condizionati.

Un esempio è suggerito dalla teoria stessa delle equazioni algebriche. Dopo aver calcolato  $x_+$ , è possibile determinare  $x_-$  non impiegando la formula esplicita, bensì sfruttando il fatto che  $x_-$  può essere calcolato come il prodotto delle radici  $x_-x_+$  diviso per  $x_+$

$$x_- = \frac{x_-x_+}{x_+} = \frac{d^2 - (d^2 - 1)}{d + \sqrt{d^2 - 1}} = \frac{1}{d + \sqrt{d^2 - 1}},$$

dove il prodotto delle radici  $x_-x_+ = d^2 - (d^2 - 1) = 1$ . Osserviamo che in questo caso non c'è nessuna cancellazione di cifre significative, perché stiamo sommando due numeri dello stesso segno.

Questo semplice esempio evidenzia come, a un problema matematicamente ben condizionato, possano corrispondere metodi numerici o algoritmi che risultano sia ben condizionati sia mal condizionati. È pertanto fondamentale prestare attenzione e privilegiare l'utilizzo di algoritmi ben condizionati.

Esaminiamo ora un quarto esempio, relativo al calcolo della radice quadrata. In particolare, consideriamo il problema di determinare la radice quadrata di un numero  $d$  positivo

$$x = \sqrt{d}.$$

Possiamo trasformare questo problema in uno dei problemi visti finora

$$x = \sqrt{d} \Leftrightarrow x^2 = d \Leftrightarrow x^2 - d = 0.$$

Questo problema  $x^2 - d = 0$  ha la struttura ancora di equazione di secondo grado. Un possibile metodo numerico che si può utilizzare per calcolare le radici quadrate è il metodo di Newton, la cui formula ricorsiva è

$$\text{dato } x_0, \quad x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}, \quad n \geq 1.$$

dove  $f(x) = x^2 - d$ . Di conseguenza, partendo ad esempio da  $x_0 = 1$ , si ha la successione di valori  $x_n$  per  $n \geq 1$

$$x_n = x_{n-1} - \frac{x_{n-1}^2 - d}{2x_{n-1}} = \frac{x_{n-1}^2 + d}{2x_{n-1}} = f_n(d).$$

In particolare, tale formula fornisce il valore  $n$ -esimo, indicato come  $f_n(d)$ . Di conseguenza, si calcola il valore  $x_n$  in funzione del dato  $d$ . Nel primo nucleo abbiamo visto che il limite di questo processo iterativo corrisponde alla soluzione del problema, ossia ovvero  $x = \sqrt{d}$

$$\lim_{n \rightarrow \infty} x_n = \sqrt{d}.$$

Quindi possiamo pensare di calcolare la radice di un numero, utilizzando il metodo di Newton. Il numero di condizionamento numerico di questo problema, applicando la formula del numero di condizionamento, risulta

$$K_n \simeq \|f'_n(d_n)\| \frac{\|d_n\|}{\|f_n(d_n)\|} = \frac{1}{2x_{n-1}} \frac{|d|}{|x_{n-1}^2 + d|} |2x_{n-1}| = \frac{|d|}{|x_{n-1}^2 + d|},$$



quindi

$$K_n \simeq \frac{|d|}{|x_{n-1}^2 + d|}.$$

Se  $n$  tende all'infinito, poiché abbiamo visto che  $x_n$  converge ad  $x$ , possiamo sostituire, nell'espressione di  $K_n$ ,  $x_{n-1}^2$  con  $x^2$

$$K_n = \frac{|d|}{|x^2 + d|},$$

ma poiché  $x = \sqrt{d}$ , abbiamo che

$$K_n = \frac{|d|}{|d + d|} = \frac{|d|}{2|d|} = \frac{1}{2}.$$

Abbiamo dunque trovato che in questo caso il numero di condizionamento numerico è uguale a  $\frac{1}{2}$  ed è indipendente da  $n$ . Inoltre,  $K_n$  è esattamente uguale al numero di condizionamento del problema matematico di partenza. Questo, dunque, è un esempio di metodo numerico stabile con un numero di condizionamento numerico pari al numero di condizionamento del problema di partenza.