

L'ALGORITMO DI STURM

Michele Impedovo, Simone Pavanelli

Lettera P.RI.ST.EM, n° 10, dicembre 1993

Questo lavoro nasce dalla collaborazione tra un insegnante e uno studente; lo studente ha curato interamente la costruzione dell'algoritmo e la realizzazione del programma (in Pascal 6) per la risoluzione del problema che qui presentiamo: la ricerca delle radici di un polinomio a coefficienti razionali.

Il problema della determinazione delle radici reali di una equazione polinomiale

$$a_n x^n + \dots + a_1 x + a_0 = 0$$

con $a_i \in \mathbb{R}$, non è in generale di semplice soluzione. Come è noto, per il teorema di Ruffini-Abel, non è possibile esprimere le soluzioni di una generica equazione polinomiale di grado n come funzioni algebriche dei coefficienti se n è maggiore di 4; inoltre per $n=3$ e $n=4$ le formule risolutive acquistano significato solo se interpretate nel campo \mathbb{C} dei numeri complessi. Infine non esiste un metodo generale per la scomposizione di un polinomio di grado n se le radici sono irrazionali: l'unica strada è quella di *approssimare* le radici, mediante opportuni algoritmi.

Anche questa strada è però insidiosa: il punto di partenza per ogni algoritmo che approssimi gli zeri di un polinomio $f(x)$ consiste nel determinare due valori a e b sull'asse reale tali che

$$f(a) \cdot f(b) < 0.$$

La continuità di $f(x)$ ci assicura allora che esiste uno zero di $f(x)$ compreso tra a e b ; tuttavia determinare due valori a e b cosiffatti non è sempre immediato: è ben vero che con un procedimento iterativo si può setacciare l'asse x partendo da un valore x_0 , e incrementare ogni volta di un dato Δx , di modo che $x_{i+1} = x_i + \Delta x$, e valutare se $f(x_i) \cdot f(x_{i+1}) < 0$; ma $f(x)$ potrebbe ammettere due zeri molto vicini, o comunque entrambi compresi tra x_i e x_{i+1} , e il passaggio da x_i a x_{i+1} li ignorerebbe entrambi. Inoltre questo metodo si rivela inefficace per trovare una radice doppia (in generale di molteplicità pari), in corrispondenza della quale $f(x)$ non cambia segno.

Il teorema di Sturm

Il Teorema di Sturm (Charles Sturm, 1803-1855, professore di Meccanica alla Sorbona di Parigi e collaboratore di Joseph Liouville) permette di determinare il numero di zeri reali di un polinomio compresi tra due numeri dati. Vogliamo qui presentare questo teorema (tralasciando la dimostrazione, che si può trovare nella bibliografia citata) e il modo con cui è stato progettato l'algoritmo e il programma relativo.

Sia dato un polinomio a coefficienti reali $f(x)$, e la sua derivata prima, $f'(x)$. Applichiamo l'*algoritmo di Euclide*, opportunamente modificato; otteniamo la cosiddetta *sequenza di Sturm*, nel seguente modo: dividiamo $f(x)$ per $f'(x)$, ottenendo come quoziente $q_1(x)$ e come resto $r_1(x)$:

$$f(x) = q_1(x) f'(x) + r_1(x).$$

Si ponga $s_1(x) = -r_1(x)$, e si divida ora per $s_1(x)$:

$$f'(x) = q_2(x) s_1(x) + r_2(x);$$

Sia $s_2(x) = -r_2(x)$, e si prosegua dividendo $s_1(x)$ per $s_2(x)$:

$$s_1(x) = q_3(x) s_2(x) + r_3(x),$$

e così via, fino a ottenere l'ultimo polinomio $s_k(x)$ non nullo. La sequenza di polinomi (di gradi decrescenti)

$$f(x), f'(x), s_1(x), s_2(x), s_3(x), \dots, s_k(x)$$

è detta *sequenza di Sturm*. Si osservi che, analogamente a quanto accade applicando l'algoritmo di Euclide in \mathbb{Z} , il polinomio $s_k(x)$ è un MCD tra $f(x)$ e $f'(x)$.

Per esempio, se $f(x) = x^4 - 3x - 1$, la sequenza di Sturm è

$$f(x) = x^4 - 3x - 1$$

$$f'(x) = 4x^3 - 3$$

$$s_1(x) = \frac{9}{4}x + 1$$

$$s_2(x) = \frac{2443}{729}$$

Per ogni numero reale c , la sequenza di Sturm individua la corrispondente sequenza di numeri reali

$$f(c), f'(c), s_1(c), s_2(c), \dots, s_k(c).$$

In questa sequenza è possibile determinare il *numero di cambiamenti di segno*: un cambiamento di segno si verifica quando un numero della sequenza è positivo, e il successivo (non nullo) è negativo, o viceversa.

Definiamo così la funzione $S(x)$ da \mathbb{R} in \mathbb{N}

$$x \xrightarrow{S} n$$

che associa ad ogni reale x_0 il numero (eventualmente zero) dei *cambiamenti di segno* nella sequenza di Sturm individuata dalla funzione $f(x)$ e valutata in x_0 .

TEOREMA (di Sturm). Dato un polinomio $f(x)$ a coefficienti in \mathbb{R} , il numero di zeri reali distinti di $f(x)$ compresi tra a e b ($a < b, f(a) \neq 0, f(b) \neq 0$) è

$$S(a) - S(b).$$

Mostriamo l'applicazione di questo teorema con il polinomio $f(x)$ di cui abbiamo prima calcolato la sequenza di Sturm.

Valutiamo le quattro funzioni della sequenza per esempio con i valori -2, -1, 0, 1, 2, 3. Otteniamo la seguente tabella:

x	-2	-1	0	1	2	3
$f(x)$	21	3	-1	-3	9	71
$f'(x)$	-35	-7	-3	1	29	105
$s_1(x)$	-7/2	-5/4	1	13/4	11/2	31/4
$s_2(x)$	2443/729	2443/729	2443/729	2443/729	2443/729	2443/729
$S(x)$	2	2	1	1	0	0

Come si può dimostrare, la funzione $S(x)$ decresce, e segnala gli intervalli nei quali è compreso uno zero di $f(x)$; non ci sono zeri tra -2 e -1, né tra 0 e 1, né tra 2 e 3, mentre uno zero è compreso tra -1 e 0, e un altro tra 1 e 2. Inoltre deduciamo che non ci sono altri zeri per $x > 2$ (infatti $S(x)$, per definizione, non può essere negativo). Potrebbero tuttavia esserci altri zeri per $x < -2$. Si pone dunque il problema di determinare un intervallo sull'asse reale che contenga *tutti* gli zeri del polinomio dato.

Limitazione degli zeri.

Il seguente teorema assicura che *tutti* gli zeri di un polinomio sono compresi tra due valori che dipendono dai coefficienti, e che possiamo facilmente calcolare.

TEOREMA. Sia $f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0$ un polinomio a coefficienti in \mathbb{R} , con coefficiente direttivo $a_n = 1$. Definiamo

$$M = \max\{|a_{n-1}|, |a_{n-2}|, \dots, |a_0|\}$$

Se c è uno zero di $f(x)$ allora

$$-(M+1) < c < M+1.$$

Osservazione. L'ipotesi $a_n = 1$ non è limitativa: dividendo il polinomio per a_n si ottiene un nuovo polinomio che ha tutti gli zeri in comune con il precedente.

Dimostrazione. Dimostriamo innanzitutto che per tutti gli $x > M+1$ risulta $f(x) > 0$ (e quindi $f(x)$ non si annulla). Infatti dalla definizione segue che per ogni $i=0,1,\dots,n-1$, $M > -a_i$, cioè $a_i > -M$; per ogni x positivo risulta allora

$$f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0 \geq -M(x^n + x^{n-1} + \dots + 1) = \\ = x^n - M \frac{x^n - 1}{x - 1} = \frac{x^n(x - (M + 1)) + M}{x - 1};$$

per qualunque $x > M+1 > 1$ risulta $f(x) > 0$.

Abbiamo così ottenuto una limitazione superiore per gli zeri di $f(x)$. Per ottenere una limitazione inferiore è sufficiente osservare che il polinomio $f(-x)$ ha radici opposte rispetto a $f(x)$, mentre i suoi coefficienti sono, in valore assoluto, gli stessi di $f(x)$, e quindi M è lo stesso; applicando a $f(-x)$ quanto abbiamo appena dimostrato, si ottiene la tesi.

Applichiamo questo teorema all'esempio precedente $f(x) = x^4 - 3x - 1$: risulta $M=3$; perciò tutti gli zeri del polinomio $x^4 - 3x - 1$ sono compresi tra -4 e 4 ; per dimostrare allora che $f(x)$ ammette solo due zeri reali distinti è sufficiente osservare che $S(-4) = 2 = S(-2)$, e per il teorema di Sturm tra -4 e -2 non sono comprese radici di $f(x)$.

La limitazione ottenuta da questo teorema, che è molto semplice da ricordare, può essere migliorata. Per esempio è possibile dimostrare (il risultato è noto come *teorema di Mac Laurin*) che è sufficiente definire M come il massimo dei valori assoluti dei coefficienti negativi per avere una limitazione superiore:

$$M = \max\{|a_j|, a < 0\}.$$

Analizziamo per esempio il polinomio

$$p(x) = x^4 - 3x^3 + 5x^2 + x + 10;$$

il teorema che abbiamo prima dimostrato afferma che tutti gli zeri sono compresi tra -11 e 11 ; il teorema di Mac Laurin ci consente invece di affermare che una radice di $p(x)$ non supera 4 ; inoltre poiché

$$p(-x) = x^4 - 3x^3 + 5x^2 + x + 10,$$

uno zero di $p(-x)$ non può superare 2 , cioè gli zeri di $p(x)$ sono maggiori o uguali a -2 ; anziché cercare gli zeri di $p(x)$ nell'intervallo $[-11, 11]$, possiamo allora limitarci al più vantaggioso intervallo $[-2, 4]$.

Un altro metodo per la limitazione degli zeri di un polinomio è noto come *teorema di Laguerre*.

TEOREMA (di Laguerre). Sia $P(x)$ un polinomio a coefficienti reali

$$x^n + a_{n-1}x^{n-1} + \dots + a_0$$

e sia $s > 0$; si divida $P(x)$ per $(x-s)$, ottenendo un quoziente $Q(x)$ e un resto R . Se tutti i coefficienti di $Q(x)$ e R risultano non negativi, allora s è una limitazione superiore per gli zeri di $P(x)$.

Dimostrazione. La dimostrazione è immediata: poiché

$$P(x) = (x-s) \cdot Q(x) + R$$

se tutti i coefficienti di $Q(x)$ sono positivi (o nulli), e $R > 0$, allora per ogni x tale che $x > s > 0$ risulta $P(x) > 0$.

Si può utilizzare questo teorema per migliorare la limitazione degli zeri, anche rispetto al teorema di Mac Laurin. Per esempio per il polinomio

$$p(x) = x^4 - 10x^3 + 20x^2 + x - 1$$

il primo teorema fornisce la limitazione $[-21, 21]$, mentre il teorema di Mac Laurin fornisce la limitazione $[-2, 11]$; applichiamo il teorema di Laguerre, con diversi valori di s :

$s=8$, coefficienti di $Q(x)$: 1, -2, 4, 33; $R=263$; un coefficiente è negativo, s è troppo piccolo;

$s=9$, coefficienti di $Q(x)$: 1, -1, 11, 100; $R=899$; c'è ancora un coefficiente negativo, ma minore in modulo;

$s=10$, coefficienti di $Q(x)$: 1, 0, 20, 201 $R=2009$.

Come si vede, tutti i coefficienti sono non negativi: 10 è la migliore maggiorazione che possa fornire il teorema di Laguerre (infatti uno dei coefficienti è nullo).

Per una limitazione inferiore, come al solito, limitiamo superiormente gli zeri di

$$p(-x) = x^4 + 10x^3 + 20x^2 - x - 1;$$

$s=1$, coefficienti di $Q(x)$: 1, 11, 31, 30; $R=29$;

tutti i coefficienti sono positivi, quindi -1 è una limitazione inferiore, ma si può migliorare:

$s=0.5$, coefficienti di $Q(x)$: 1, 10.5, 25.25, 11.625; $R=4.813$;

$s=0.25$, coefficienti di $Q(x)$: 1, 10.25, 22.6, 4.6; $R=0.16$;

$s=0.2$, coefficienti di $Q(x)$: 1, 10.2, 22.04, 3.4; $R=-0.32$;

Come si vede -0.25 è una limitazione inferiore, ma non -0.2 ; riassumendo, con il metodo di Laguerre abbiamo per gli zeri di $p(x)$ la limitazione $[-0.25, 10]$. Poiché il calcolo dei coefficienti di $Q(x)$ e di R è molto semplice dal punto di vista algoritmico (si simula la cosiddetta *regola di Ruffini*), il metodo di Laguerre si presta ad essere utilizzato con profitto in un programma.

La molteplicità degli zeri .

Il teorema di Sturm fornisce anche informazioni sulla molteplicità degli zeri del polinomio $f(x)$.

Osserviamo che se uno zero di $f(x)$, $x=c$, è di molteplicità maggiore di 1, allora

$$f(x) = (x-c)^2 g(x)$$

$$f'(x) = 2(x-c)g(x) + (x-c)g'(x) = (x-c)[2g(x) + (x-c)g'(x)],$$

cioè c è uno zero anche di $f'(x)$; in questo caso l'ultimo polinomio $s_k(x)$ della sequenza di Sturm, che, ricordiamo, è un MCD tra $f(x)$ e $f'(x)$, non è una costante, bensì un polinomio almeno di primo grado, e i suoi zeri coincidono con gli zeri di $f(x)$ di molteplicità maggiore di 1. Allora, se $s_k(x)$ è di 1° grado $f(x)$ ammette uno zero doppio, se $s_k(x)$ è di 2° grado $f(x)$ ammette due zeri doppi (oppure uno zero triplo), e così via.

Nell'esempio $f(x)=x^4 - 3x - 1$, l'ultimo polinomio della sequenza di Sturm è la costante $2443/729$, perciò $f(x)$ e $f'(x)$ sono primi tra loro, e $f(x)$ ammette solo zeri semplici. Possiamo concludere allora che $f(x)$ ammette due zeri reali, $-1 < x < 0$, $1 < x < 2$, e due zeri complessi coniugati.

Il programma (in Turbo Pascal 6).

Il programma chiede in ingresso il grado n del polinomio e gli $n+1$ coefficienti, e fornisce in uscita:

- 1) la sequenza di Sturm;
- 2) gli estremi di un intervallo che contiene tutte le radici reali;
- 3) la valutazione della funzione di Sturm agli estremi di tale intervallo e il numero di zeri reali distinti;
- 4) le radici approssimate.

Il programma richiede in ingresso coefficienti interi; questa scelta radicale (che ovviamente riguarda soltanto il programma implementato, e non il teorema di Sturm) è motivata dal fatto che usando coefficienti di tipo *real*, nei numerosi calcoli necessari a costruire la sequenza di Sturm si avrebbe una inevitabile e massiccia propagazione degli errori.

Nel programma viene definito il tipo FRAZIONI come un *record* a due campi (numeratore e denominatore) di *integer*. Viene quindi definito il tipo POLINOMI come un *record* a due campi: il campo GRADO è un *integer* e contiene n , il campo COEF è un vettore (*array*) di FRAZIONI, cioè di coppie ordinate di *integer* (inizialmente i denominatori sono tutti posti uguali a 1); infine si definisce il tipo SEQUENZA, che è un vettore di POLINOMI; il primo polinomio è quello dato in ingresso, il secondo è la sua derivata, che viene calcolata dal programma.

Il programma calcola il resto della divisione tra $f(x)$ e $f'(x)$ (le operazioni vengono effettuate tra variabili di tipo FRAZIONI, quindi sono definite tutte le operazioni e la semplificazione di frazioni tramite opportune *function* e *procedure*), cambia segno ai coefficienti del resto (in realtà cambia segno ai numeratori, poiché in tutto il programma si fa in modo che i denominatori restino positivi), e li assegna alla terza componente della variabile di tipo SEQUENZA, e così via, fino ad ottenere l'ultimo polinomio non nullo $s_k(x)$.

Vengono quindi scritti sullo schermo i polinomi della sequenza di Sturm.

Dopodiché si procede alla limitazione degli zeri; il programma calcola $M=\max\{|a_i|\}$ e poi provvede a migliorare l'intervallo $[-(M+1), M+1]$ mediante il metodo di Laguerre, determinando così l'intervallo $[inf, sup]$.

Partendo da $c=inf$ si calcola la funzione di Sturm $S(c)$, si incrementa c di $\Delta x=1$, $c'=c+1$, fino a che $S(c)>S(c')$. Se $S(c)-S(c')=1$ allora in $[c, c']$ è compresa una radice di $f(x)$; se invece $S(c)-S(c')>1$ (significa che tra c e c' sono comprese più radici), si diminuisce Δx , fino a trovare un c' tale che $S(c)-S(c')=1$.

A questo punto sappiamo che tra c e c' è compresa una sola radice del polinomio. Tra i tanti metodi di approssimazione abbiamo infine optato per la soluzione più facile e più efficace, quella di utilizzare ancora la funzione di Sturm: infatti occorre semplicemente cercare, tra c e c' , due valori a e b (tanto più vicini quanto maggiore è la precisione voluta) tali che

$S(a)-S(b)=1$. Poiché i coefficienti dei polinomi della sequenza di Sturm sono, nel programma, le coppie ordinate di interi che definiscono le FRAZIONI, non si corre alcun rischio di propagazione indesiderata dell'errore: tutti i risultati parziali sono esatti. Con questo metodo, a differenza di altri metodi di approssimazione, non si cerca il punto in cui $f(x)$ cambia segno (cosa che non accade per zeri di molteplicità di ordine pari), bensì il punto in cui la funzione di Sturm $S(x)$ diminuisce di una unità.

Michele Impedovo

docente presso il Liceo Classico di S. Donato M.se

Simone Pavanelli

studente della classe II liceo classico sez. B.

Bibliografia.

1. L. Childs: *Algebra, un'introduzione concreta*, ETS editrice.
2. R. Beaumont, R. Pierce: *The algebraic foundation of Mathematics*, Addison-Wesley.
3. A. Lissoni, S. Salsa: *Matematica e laboratorio*, Ghisetti e Corvi.