INCONTRI CON LA MATEMATICA N. 14

Castel San Pietro Terme, 3-5 novembre 2000

Datemi un polinomio e vi solleverò il mondo: strutture e approssimazioni

prof. Michele Impedovo
Liceo Scientifico Galileo Ferraris di Varese

Il concetto di rigore si è modificato nella storia della matematica:
le successive presentazioni costituiscono una successione;
si tratta di una successione monotona? o convergente?.
Ho forti dubbi in entrambi i casi.
Modesto Dedò

Introduzione

Si può dire che il tema "polinomi" accompagni costantemente il curriculum di matematica in tutto il ciclo secondario: dai prodotti notevoli fino ai polinomi di Taylor.
Tuttavia sembrano non emergere, nella nostra programmazione, due aspetti che considero centrali, e che potrebbero costituire un paradigma generale per un mutamento di rotta nell'insegnamento della matematica nel 2000: minor attenzione alle tecniche di calcolo e maggior attenzione ai concetti, agli algoritmi, alla produzione di congetture, agli esperimenti. I due aspetti di cui parlo sono i seguenti:

- la struttura algebrica dei polinomi: impadronirsi dei (pochi) concetti strutturali potrebbe evitare lunghe litanie sintattiche (spesso imprecise o addirittura sbagliate) e percorsi ad ostacoli nello sviluppo del calcolo letterale.
- il polinomio è un potente strumento di approssimazione: ora che possiamo agevolmente utilizzare strumenti di calcolo automatico e ora che la *computer algebra* è disponibile su calcolatrici di piccole dimensioni, gli algoritmi di approssimazione (che sono stati spesso motore di sviluppo nella storia della matematica) sono alla portata di uno studente secondario, e costituiscono un modo efficace (e non noioso) di approccio al calcolo infinitesimale.

Negli esempi e nei calcoli proposti è stata usata la calcolatrice TI-92, nell'ambito della sperimentazione LABCLASS (Laboratorio in classe) promossa dalla Direzione Classica del MPI.

Struttura

Il calcolo letterale, come insegnava Giovanni Prodi ormai quasi trent'anni fa, dovrebbe iniziare dai polinomi, e non dai monomi. La definizione di monomio infatti è difficilmente strutturabile e si presta ad equivoci che puntualmente compaiono sui nostri libri di testo più pedanti (si provi a stabilire se, in base alle varie definizioni, $\left(\sqrt{2}+\ln 2\right)x$ è o non è un monomio, oppure si provi a spiegare che cosa significhi esattamente che "due monomi si possono sommare solo se sono simili"). Si tenga anche conto del fatto che l'approccio al calcolo letterale viene solitamente svolto alle scuole medie (pur non essendo previsto dai vigenti programmi) da non-matematici, che, nella loro esagerata paura del rigore, finiscono per credere a questa cattiva interpretazione della matematica.

Tutte le cosiddette "regole" del calcolo letterale sono da ricondursi essenzialmente alle caratteristiche della struttura algebrica di **campo**, e quindi sono tutte riassumibili in poche, semplici proprietà che riguardano le due operazioni fondamentali di addizione e moltiplicazione.
Scegliamo un *campo* numerico, per esempio il campo **Q** dei razionali oppure il campo **R** dei reali. All'insieme **Q** (oppure **R**) aggiungiamo un simbolo, per esempio "x" (che non è un numero: è una variabile, un contenitore).
Mescoliamo il tutto (con addizioni e moltiplicazioni, chiedendo che valgano le stesse proprietà che valgono in **Q**) e quello che ne esce è l'*anello dei polinomi in x*, che indichiamo con **R**$[x]$: si tratta di una struttura fondamentale, in algebra e in geometria (si pensi ai luoghi geometrici caratterizzati da un'equazione algebrica).
Da un punto di vista algoritmico un polinomio di grado n è univocamente determinato dalla lista dei suoi n+1 coefficienti. In effetti potremmo abbreviare la scrittura

$$a_n x^n + a_{n-1} + \ldots + a_1 x + a_0$$

semplicemente con la lista

$$(a_0, a_1, \ldots, a_{n-1}, a_n)$$

(è esattamente quello che si fa quando si scrive un numero naturale: la variabile x è la base del sistema di numerazione). Così possiamo definire con esattezza le operazioni di addizione e moltiplicazione tra due polinomi.

$a(x)=(a_0, a_1, \ldots, a_{n-1}, a_n, \ldots)$

$b(x)=(b_0, b_1, \ldots, b_{m-1}, b_m, \ldots)$

$a(x)+b(x) = (a_0+b_0, a_1+b_1, a_2+b_2, \ldots\ )$

$a(x)\cdot b(x) = (a_0b_0, a_1b_0+a_0b_1, a_2b_0 + a_1b_1 + a_0b_2, \ldots, \sum_{k=0}^{n} a_{n-k}b_k\ , \ldots)$

Esistono interessanti analogie tra l'anello degli interi **Z** e l'anello dei polinomi su un campo A. Come in **Z**, anche in $A[x]$ è possibile definire le operazioni "**div**" e "**mod**", cioè è possibile definire una *divisione con resto*, e quindi applicare l'algoritmo euclideo (per esempio per il calcolo del MCD tra polinomi).
Così come in **Z** è possibile parlare di numeri *primi* e numeri *composti*, anche nell'anello dei polinomi su un campo A è possibile dare una definizione di polinomi *primi* (*irriducibili*) e *composti* (*riducibili*). Il problema della fattorizzazione di un polinomio andrebbe affrontato solamente dopo aver precisato il campo numerico al quale si fa riferimento. Nulla vieta (anzi, dal punto di vista didattico è assai interessante) considerare il campo finito $\mathbf{Z}_p$ e l'anello dei polinomi $\mathbf{Z}_p[x]$.
Succede che ad ogni polinomio $p(x)\in A[x]$ sia associata in modo naturale una funzione da A in A:

$$x \to p(x).$$

Nella pratica didattica si rischia spesso di incorrere in un equivoco: quello di confondere l'aspetto sintattico (il polinomio) con l'aspetto semantico (la funzione). Sono oggetti differenti, e come tali andrebbero trattati. Perché confondere le giovani menti dicendo che l'uguaglianza

$$\frac{x^2-1}{x+1} = x-1$$

"non è sempre vera" (quindi è falsa) quando è evidente che

$$(x+1)(x-1) = x^2-1$$

è vero?
Si osservi che due polinomi $f(x)$ e $g(x)$ sono uguali se hanno lo stesso grado e gli stessi coefficienti, cioè se sono *sintatticamente* uguali, cioè ancora se sono lo stesso elemento di $A[x]$. Due funzioni, f e g, sono uguali se hanno lo stesso dominio D, e per ogni $a\in D$ risulta $f(a)=g(a)$. È ovvio che due polinomi uguali generano due funzioni uguali. Non è affatto ovvio il viceversa: è possibile che due polinomi distinti $f(x)$, $g(x)\in A[x]$ generino la stessa funzione, cioè per ogni $a\in A$ risulti $f(a)=g(a)$?

A questa domanda risponde negativamente il *teorema di identità dei polinomi*, che (fatto poco noto) richiede una ipotesi: il campo A deve essere *infinito*. Solo in questo caso due polinomi sono uguali se e solo se lo sono le corrispondenti funzioni polinomiali. Controesempio: in $\mathbf{Z}_5[x]$ i polinomi

$$f(x) = x^5+x^3-x+1, \qquad g(x) = x^3+1.$$

$f(x)$ e $g(x)$ sono polinomi diversi. Tuttavia

$$f(0) = g(0) = 1$$
$$f(1) = g(1) = 2$$
$$f(2) = g(2) = 4$$
$$f(3) = g(3) = 3$$
$$f(4) = g(4) = 0$$

cioè $f(x)$ e $g(x)$ sono uguali come funzioni.

È il *teorema di Ruffini* a far da ponte tra l'aspetto sintattico (il polinomio $f(x)$) e l'aspetto semantico (la funzione che esso rappresenta): $f(a)=0$ (aspetto semantico) se e solo se $f(x)=(x-a)g(x)$ (aspetto sintattico).
Di questo fondamentale teorema, a dir la verità, resta ben poca cosa nelle aule scolastiche: si preferisce annegarlo con la famigerata *regola di Ruffini*, con tecniche di calcolo oscure e mal interpretate dai più.
L'analisi di questa struttura algebrica conduce in modo naturale al problema della risolubilità delle equazioni algebriche e, in prospettiva, alla necessità di costruire un campo *algebricamente chiuso* (cioè che contiene tutte le radici di un polinomio a coefficienti nel campo stesso), alla definizione quindi del campo **C** dei numeri complessi, e infine al *teorema fondamentale dell'algebra*.

Approssimazioni

Lo slogan potrebbe essere il seguente:

"Si può fare tutto con addizioni e moltiplicazioni".

Questa è una caratteristica preziosa dei polinomi: il fatto che addizioni e moltiplicazioni siano facilmente implementabili rende i polinomi insostituibili nel calcolo automatico. Come calcolare i valori di una funzione trascendente $f(x)$ in un punto x_0? La nostra calcolatrice ha n cifre decimali: se troviamo un polinomio $p(x)$ tale che

$$\left| f(x_0) - p(x_0) \right| < 0.5 \cdot 10^{-n}$$

il gioco è fatto.
Vale la pena di citare subito due risultati fondamentali, che danno senso e struttura al problema.
Teorema di Weierstrass (Karl Weierstrass 1815-1897). Data una funzione $f(x)$, continua su un intervallo $[a, b]$, per ogni $\varepsilon>0$ esiste un polinomio $p(x)$ tale che

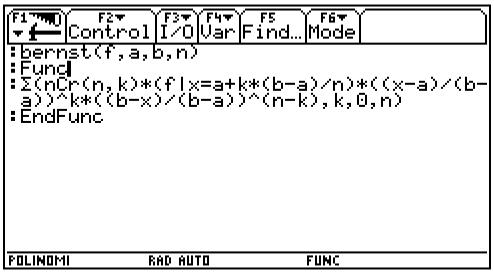
$$\left| f(x) - p(x) \right| < \varepsilon$$

per ogni $x \in [a, b]$.
Questo teorema è potentissimo: la sola ipotesi di continuità garantisce l'esistenza di un polinomio che è vicino a f quanto si vuole.
Il secondo risultato è complementare al primo: si tratta di un teorema costruttivo, che esibisce una successione di polinomi che converge a f, e la cui dimostrazione è quindi implicitamente una dimostrazione del teorema di Weierstrass.
Teorema di Bernstein (Sergi Natanovich Bernstein 1880-1968). La successione di polinomi (*polinomi di Bernstein*)

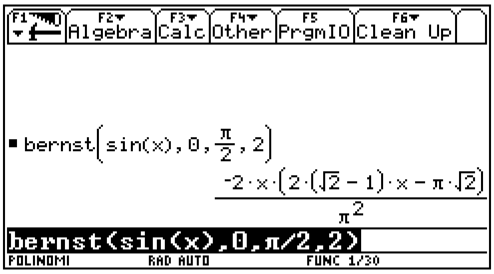
$$B_n(x) := \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}$$

converge a f su $[a, b]$, nel senso che per ogni $\varepsilon>0$ esiste un naturale N tale che per ogni $n>N$ e per ogni $x\in[a, b]$ risulta $\left|f(x)-B_n(x)\right|<\varepsilon$.

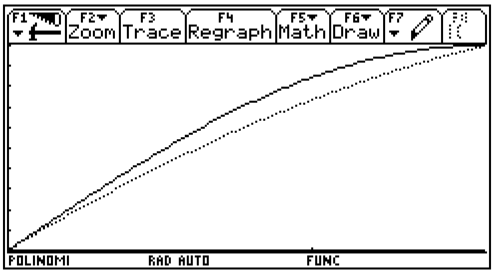
Il programma bernstein(f,a,b,n) fornisce il polinomio di grado n della successione.



Vediamo per esempio il polinomio di secondo grado che approssima $\sin(x)$ su $[0,\pi/2]$.



Confrontiamo il grafico di tale polinomio (punteggiato) con quello di $\sin(x)$ nel rettangolo $[0,\pi/2]\times[0,1]$.



Come si vede l'approssimazione è modesta; ci servono algoritmi di approssimazione più efficienti.
Possiamo lavorare per **interpolazione**: se conosciamo i valori che $f(x)$ assume in certi punti

$$x_0, x_1, \ldots, x_n$$

allora possiamo determinare il polinomio di grado n che li soddisfa, e ci aspettiamo che quel polinomio sia, relativamente ad un certo intervallo, una buona approssimazione di $f(x)$.
Possiamo lavorare per **approssimazione locale**: se conosciamo, relativamente ad un punto x_0, i valori $f(x_0)$, $f'(x_0)$, $f''(x_0)$, …, $f^{(n)}(x_0)$ allora possiamo affidarci al polinomio di Taylor di grado n di $f(x)$ con centro in x_0.
Possiamo lavorare per **approssimazione globale**, ed è questa la strada in un certo senso più interessante: data $f(x)$ continua in $[a, b]$, cerchiamo un polinomio $p(x)$ tale che per ogni $x\in[a, b]$ risulti

$$\left|f(x)-p(x)\right|<\varepsilon,$$

dove ε è il massimo errore che siamo disposti a tollerare.
Nel seguito si esemplificano tali diversi approcci per la funzione trascendente $\sin(x)$ nell'intervallo $[0, \pi/2]$.

Interpolazione

Il problema più generale è il seguente: dati $n+1$ punti

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$$

determinare il polinomio $p(x)$ di grado (al più) n che “passa” per i punti, cioè tale che

$$\begin{aligned} & p(x_0) = y_0 \\ & p(x_1) = y_1 \\ & \ldots \\ & p(x_n) = y_n \end{aligned}.$$

Un notevole teorema ci assicura del fatto che se le ascisse dei punti sono tutte distinte, allora tale polinomio <u>esiste ed è unico</u>.

Gli n coefficienti costituiscono la soluzione del sistema lineare:

$$\begin{cases} a_n x_0^n + \ldots + a_1 x_0 + a_0 = y_0 \\ a_n x_1^n + \ldots + a_1 x_1 + a_0 = y_1 \\ \ldots \\ a_n x_n^n + \ldots + a_1 x_n + a_0 = y_n \end{cases}$$

In notazione matriciale:

$$\begin{bmatrix} x_0^n & \ldots & x_0 & 1 \\ x_1^n & \ldots & x_1 & 1 \\ & \ddots & & \\ x_n^n & \ldots & x_n & 1 \end{bmatrix} \begin{bmatrix} a_n \\ \ldots \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} y_0 \\ \ldots \\ y_{n-1} \\ y_n \end{bmatrix}$$

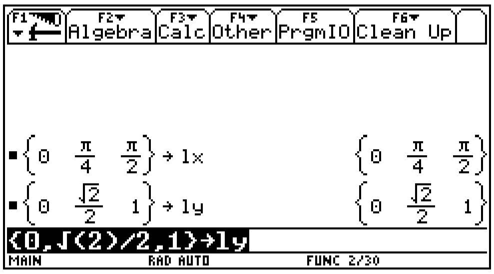
Vediamo come risolvere il problema con la TI-92.

L’obiettivo è quello di definire una funzione che prenda in ingresso due liste: la lista *lx* delle ascisse e la lista *ly* delle ordinate, e fornisca in uscita il polinomio richiesto.

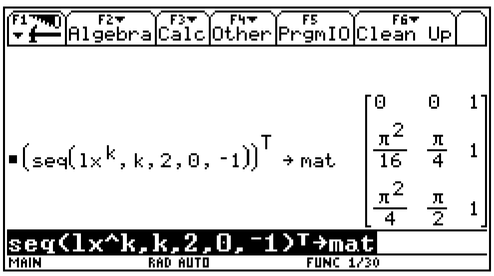
Illustriamo passo-passo il procedimento mediante un semplice esempio: vogliamo approssimare sin(x) sull’intervallo $[0, \pi/2]$, conoscendone i valori in 0, $\pi/4$, $\pi/2$. Vogliamo dunque calcolare il polinomio di secondo grado che interpola i tre punti

$$(0, 0), \left(\frac{\pi}{4}, \frac{\sqrt{2}}{2}\right), \left(\frac{\pi}{2}, 1\right).$$

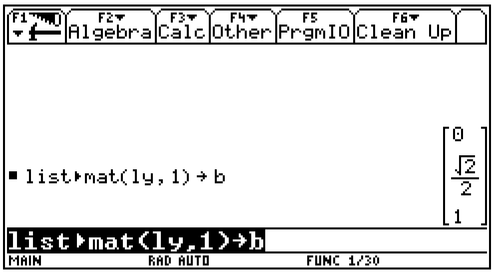
Definiamo innanzitutto le liste delle ascisse e delle ordinate.



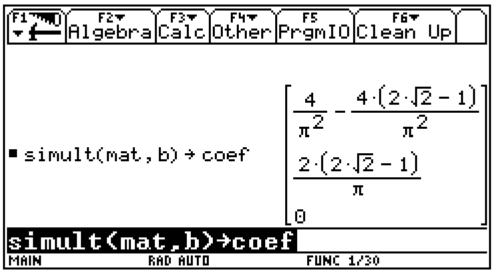
Costruiamo ora la matrice dei coefficienti. Il comando *seq*, del tutto analogo al *vector* di DERIVE, consente di costruire, per righe, la matrice delle successive potenze delle ascisse. La matrice dei coefficienti è la *trasposta* di quest’ultima.



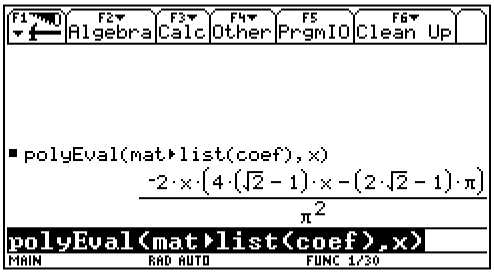
Costruiamo ora il vettore colonna dei termini noti.



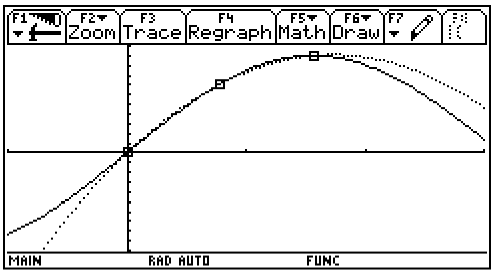
Finalmente risolviamo il sistema mediante il comando *simult*:

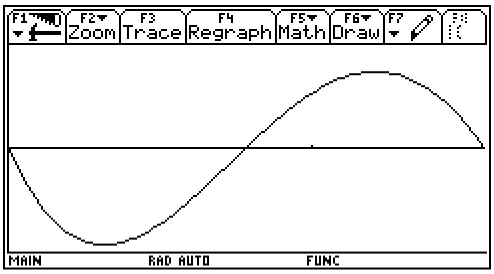


Trasformiamo i coefficienti in lista e calcoliamo infine il polinomio $p(x)$ mediante il comando *polyeval*:

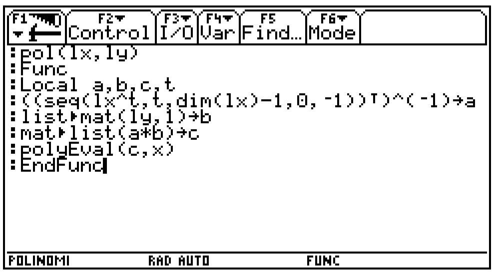


Ecco il grafico di $\sin(x)$ e $p(x)$ nel rettangolo $[-1, 3]\times[-1, 1]$ e il grafico di $\sin(x)-p(x)$ (la cosiddetta *curva d'errore*) nel rettangolo $[0, \pi/2]\times[-0.025, 0.025]$. L'errore massimo è circa 0.0235.



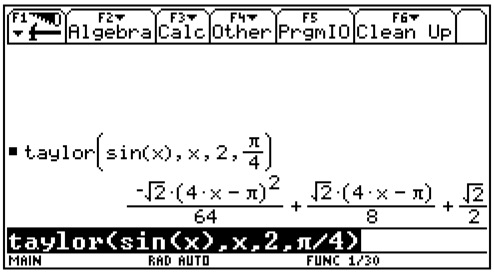


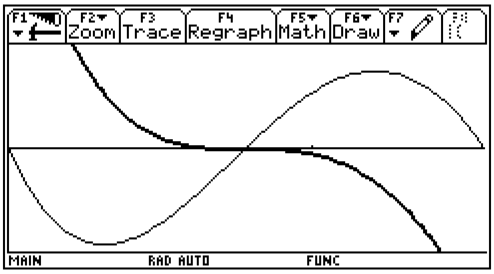
Il programma che implementa il procedimento ora visto non fa che generalizzare i vari passaggi.



Approssimazione locale

Qualunque CAS (e anche la TI-92) calcola il polinomio di Taylor di grado n con centro in x_0 di funzioni derivabili n volte. Calcoliamo il polinomio di Taylor di 2° grado di $\sin(x)$ con centro in $\pi/4$ e confrontiamo la relativa curva d'errore con quella del polinomio interpolatore calcolato precedentemente, sempre nel rettangolo $[0, \pi/2]\times[-0.025, 0.025]$.





Il grafico è illuminante: il polinomio di Taylor si allontana rapidamente da $\sin(x)$ appena ci si allontani da $\pi/4$.

Approssimazione globale

Qual è il "miglior" polinomio $p(x)$ di grado n che approssima una funzione continua $f(x)$ su un intervallo $[a,b]$? Naturalmente dobbiamo dare la definizione di "miglior polinomio", e abbiamo diverse possibilità (addirittura infinite), a seconda di quale strumento adottiamo per misurare la "distanza" tra due funzioni.

Un modo molto naturale di procedere potrebbe essere quello di misurare la distanza tra due funzioni $f(x)$ e $p(x)$ su $[a,b]$ mediante l'integrale del valore assoluto della differenza:

$$d_1(f,p) := \int_a^b |f(x) - p(x)|\,dx$$

Un altro modo di misurare la distanza tra due funzioni su un intervallo $[a, b]$ è quello di trasportare nel continuo il metodo *dei minimi quadrati*, cioè considerare, al posto della <u>somma</u> dei quadrati degli scarti da n punti (x_i, y_i), l'<u>integrale</u> del quadrato di $f(x)-p(x)$ su $[a, b]$

$$d_2(f,p) := \sqrt{\int_a^b (f(x) - p(x))^2\,dx}\,,$$

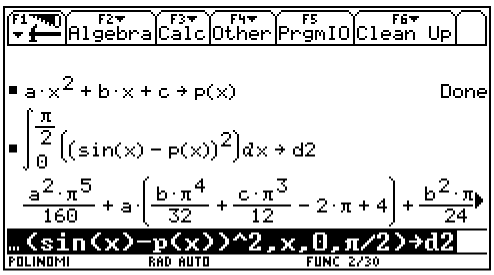
e scegliere come miglior polinomio di grado n quello che minimizza tale integrale.

Una terza scelta per definire la distanza tra due funzioni è quella già vista nei teoremi di Weierstrass e Bernstein: si chiede che $f(x)-p(x)$, per ogni $x\in[a, b]$, non superi, in valore assoluto, un certo errore. Si definisce così la distanza tra f e p nel seguente modo:

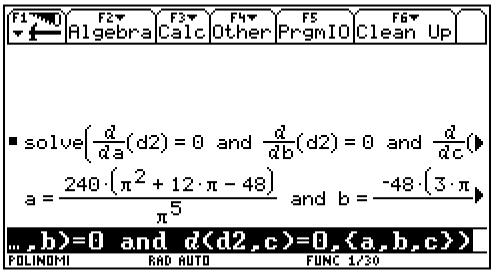
$$d_{\max}(f, p) := \max_{x\in[a,b]}\{|f(x) - p(x)|\}$$

Ciascuna distanza porta ad algoritmi differenti. Vediamo per esempio come minimizzare d_2 per $f(x)=\sin(x)$ su $[0, \pi/2]$ con un polinomio di secondo grado.

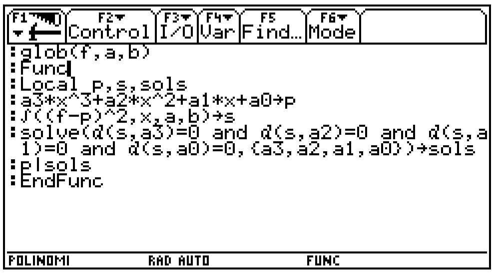
Calcoliamo $d_2(f,p)$ per $p(x)=ax^2+bx+c$, memorizzando il risultato nella variabile d2.



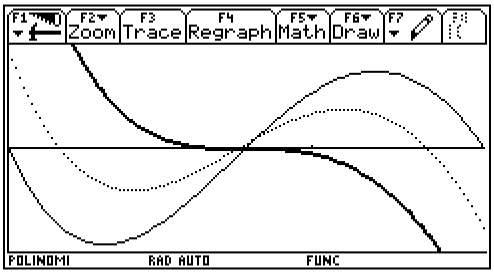
Per minimizzare d2 imponiamo che si annullino le derivate parziali rispetto ad *a*, *b*, *c*.



L’intero procedimento si può automatizzare mediante il seguente programma glob.



È evidente che si tratta di calcoli troppo laboriosi per essere svolti con carta e penna: questa attività ha senso soltanto se si può disporre di un CAS. Otteniamo dunque un polinomio che approssima molto bene sin(x) su $[0, \pi/2]$, tanto che i due grafici sono quasi indistinguibili nell’intervallo $[0, \pi/2]$. Confrontiamo la curva d’errore (tratteggiata) di quest’ultimo polinomio con quelle dei due polinomi calcolati precedentemente.



Una parziale conclusione è la seguente: nell’insegnamento della matematica del 2000 occorre dimenticare quel malinteso senso di “rigore” che tanto ha infastidito le giovani menti dei nostri allievi (anche i migliori). Occorre affidare alla <u>semantica</u> il valore del nostro insegnamento (gli algoritmi e le approssimazioni possono svolgere egregiamente questo ruolo); e quando si fa sintassi, quando si decide finalmente di formalizzare un oggetto matematico, perbacco, allora occorre fare sul serio. Altro che “due monomi si possono sommare solo se sono simili”!