1. *La Convessità*

Con le definizioni e gli strumenti che abbiamo introdotto fino a questo punto siamo in grado di distinguere una funzione il cui grafico sia del tipo illustrato in figura 1.1(a) da una il cui grafico sia quello illustrato nella figura 1.1(b)

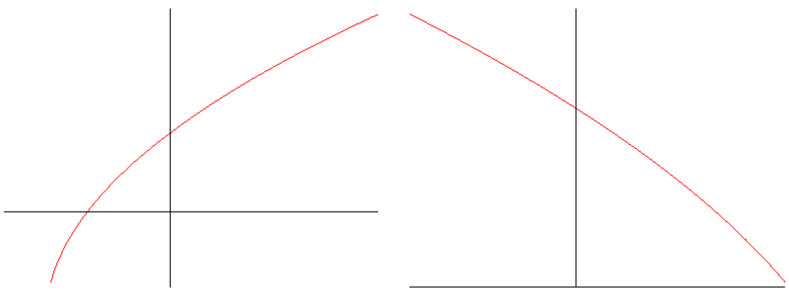

(a) (b)

Figure 1.1:

Possiamo infatti osservare che il primo è il grafico di una funzione crescente mentre il secondo rappresenta una funzione decrescente.

Abbiamo inoltre già sviluppato strumenti (studio del segno della derivata prima) che ci consentono di stabilire se una funzione è crescente o decrescente.

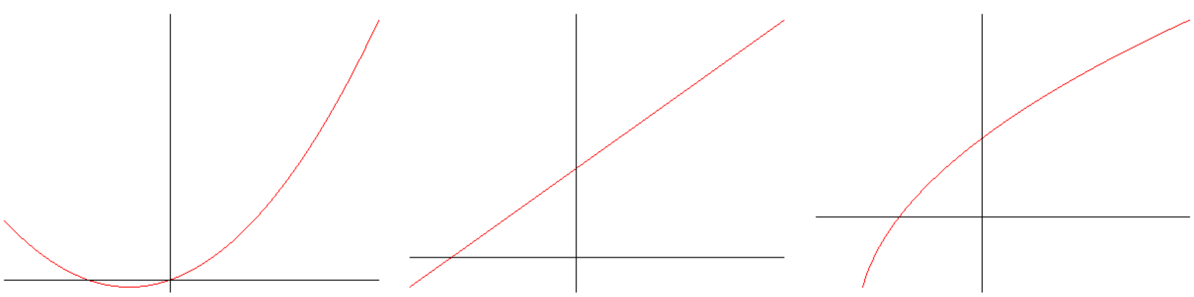

(a) (b) (c)

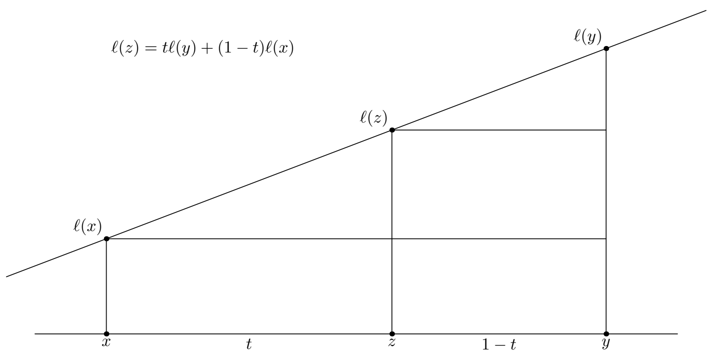
Figure 1.2:

Non siamo tuttavia ancora in grado di distinguere tra i grafici delle tre funzioni rappresentate in 1.2(a),1.2(b),1.2(c).

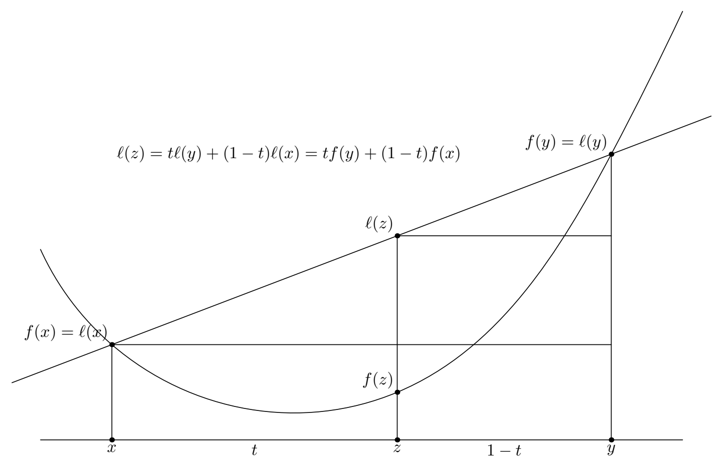
in quanto, ad un primo esame, possiamo osservare che tutte e tre sono funzioni crescenti; è tuttavia chiaro che si tratta di funzioni il cui grafico presenta caratteristiche molto diverse, così come è evidente quale è la differenza tra una scodella ed un ombrello.

Onde cercare di definire una proprietà che ci consenta di distinguere tra i tre grafici cominciamo ad esaminare il più semplice dei tre cioè il secondo. Chiaramente si tratta di una retta e quindi il suo grafico è individuato da due punti.

Indichiamo con ℓ la funzione relativa e con $(x, \ell(x))$, $(y, \ell(y))$ due punti del suo grafico. Possiamo individuare il valore di ℓ in z semplicemente usando la proporzionalità tra i triangoli indicati in figura 1.3(a).



(a)



(b)

Figure 1.3:

Avremo infatti che

$$\frac{\ell(z) - \ell(x)}{z - x} = \frac{\ell(y) - \ell(x)}{y - x} \tag{1.1}$$

Poichè

$$\frac{\ell(z) - \ell(x)}{z - x} = \frac{\ell(x) - \ell(z)}{x - z} \quad , \quad \frac{\ell(x) - \ell(y)}{x - y} = \frac{\ell(y) - \ell(x)}{y - x}$$

la 1.1 non cambia anche nel caso in cui z non sia, come in figura, interno all'intervallo di estremi x ed y. Inoltre non è restrittivo considerare $x < y$.

Avremo pertanto che il valore di ℓ in z è dato da

$$\ell(z) = \ell(x) + (z - x)\frac{\ell(y) - \ell(x)}{y - x} \tag{1.2}$$

La 1.2 è semplicemente l'equazione di una retta che passa per il punto $(x, \ell(x)$ ed ha coefficiente angolare $\frac{\ell(y)-\ell(x)}{y-x}$.

È utile osservare che, se poniamo

$$t = \frac{z - x}{y - x}$$

esprimiamo, nel contempo, la proporzionalità

$$\frac{t}{1} = \frac{z-x}{y-x}$$

tra le lunghezze dei segmenti di $[x, z]$ e $[x, y]$ ed i valori t ed 1.

Pertanto il rapporto tra le lunghezze dei segmenti $[z, y]$ e $[x, y]$, sarà uguale a $1 - t$.

Un semplice calcolo mostra infatti che

$$1 - t = 1 - \frac{z-x}{y-x} = \frac{y-x-z+x}{y-x} = \frac{y-z}{y-x}$$

Inoltre se poniamo

$$t = \frac{z-x}{y-x} \tag{1.3}$$

avremo

$$z - x = t(y - x) \tag{1.4}$$

e quindi

$$z = x + t(y-x)x = ty + (1-t)x \tag{1.5}$$

Per $t \in (0, 1)$ la 1.5 individua un punto z che si trova all'interno dell'intervallo di estremi x ed y, mentre per $t > 1$ si hanno punti a destra di y e per $t < 0$ si hanno punti a sinistra di x.

Similmente possiamo scrivere la 1.2 come

$$\begin{aligned}\ell(z) = \ell(x) + (z-x)\frac{\ell(y)-\ell(x)}{y-x} = \ell(x) + (\ell(y)-\ell(x))\frac{z-x}{y-x} \\ \ell(x) + t(\ell(y)-\ell(x)) = t\ell(y) + (1-t)\ell(x)\end{aligned} \tag{1.6}$$

ed infine possiamo scrivere

$$\ell(ty + (1-t)x) = t\ell(y) + (1-t)\ell(x) \tag{1.7}$$

ed osservare che al variare di t la 1.7 consente di esprimere il fatto che tutti i valori $\ell(z) = \ell(ty + (1-t)x)$ si trovano sulla retta di cui abbiamo studiato il grafico.

Se ora sovrapponiamo i primi due grafici nelle figure 1.3(a) e1.3(b), risulta evidente che, se chiamiamo f la funzione del primo grafico ed x e y i punti di intersezione tra il grafico e la retta, avremo che, all'interno dell'intervallo $[x, y]$, il grafico di f sta sotto il grafico della retta.

Chiamiamo una tale funzione **convessa** ed esprimiamo il fatto che abbiamo appena individuato semplicemente chiedendo che

$$f(ty + (1-t)x) \le tf(y) + (1-t)f(x) \qquad \forall t \in (0,1)$$

Poniamo in altre parole la seguente definizione

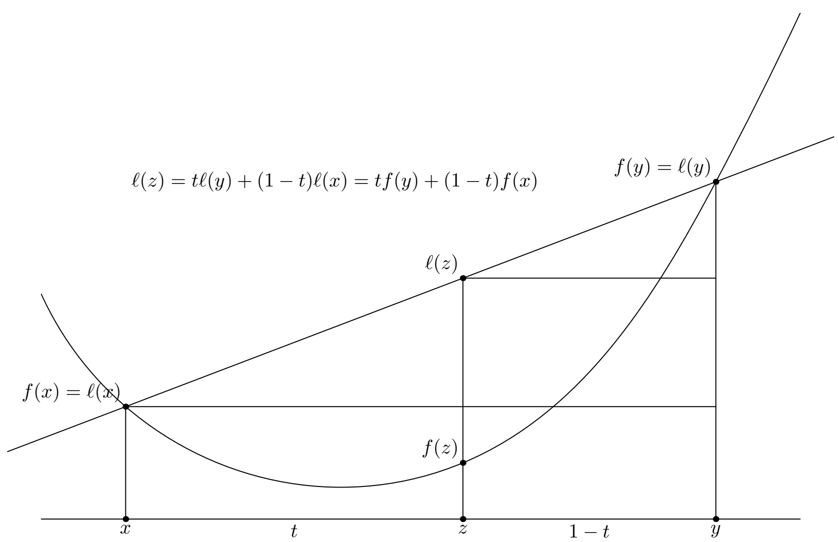
Definizione 1.1 *f si dice convessa in (a,b) se*

$$\forall x, y \in (a,b)\ ,\ \forall \lambda \in (0,1)$$

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y) \tag{1.8}$$

Diciamo che f è strettamente convessa se la 1.8 vale in senso stretto.

Figure 1.4:



Teorema 1.1 *f è convessa se e solo se, $\forall n \in N$, $\forall x_i \in (a,b)$, $\forall \lambda_i \ge 0$*

$$\sum_{i=1}^{n} \lambda_i = 1 \quad \text{si ha} \quad f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \le \sum_{i=1}^{n} \lambda_i f(x_i)$$

Dimostrazione. Proviamolo per induzione: per $n = 2$ è ovvio. Supponiamolo vero per n e proviamolo per $n+1$; sia

$$\Lambda = \sum_{i=1}^{n} \lambda_i \ = \ 1 - \lambda_{n+1},$$

si ha

$$\begin{aligned} f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) &= f\left(\Lambda \sum_{i=1}^{n} \frac{\lambda_i x_i}{\Lambda} + (1-\Lambda)x_{n+1}\right) \leq \\ &\leq \Lambda f\left(\sum_{i=1}^{n} \frac{\lambda_i}{\Lambda} x_i\right) + (1-\Lambda) f(x_{n+1}) \leq \\ &\leq \sum_{i=1}^{n} \lambda_i f(x_i) + \lambda_{n+1} f(x_{n+1}) \end{aligned}$$

□

Definizione 1.2 *Sia $A \subset \mathbb{R}^2$, diciamo che A è un insieme convesso se, per ogni $P, Q \in A$, il segmento di estremi P e Q è interamente contenuto in A.*

Osserviamo che, se

$$P = (p_1, p_2) \quad , \quad Q = (q_1, q_2)$$

un qualunque punto R del segmento di estremi P e Q ha coordinate

$$(\lambda p_1 + (1-\lambda)q_1, \lambda p_2 + (1-\lambda)q_2) \ \ , \ \ \lambda \in [0,1] \ .$$

Definizione 1.3 *Definiamo epigrafico di f l'insieme*

$$\mathrm{epi} f = \{(x, \alpha) \in \mathbb{R}^2 \ : \ f(x) \leq \alpha \ \}.$$

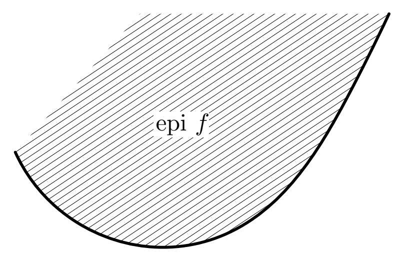


Figure 1.5: Epigrafico

Teorema 1.2 *f è convessa se e solo se $\mathrm{epi} f$ è convesso.*

Dimostrazione. Supponiamo che f sia convessa e siano

$$(x, \alpha), (y, \beta) \in \mathrm{epi} f \ \ ;$$

allora

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \leq \lambda\alpha + (1-\lambda)\beta$$

per cui

$$(\lambda x + (1-\lambda)y, \lambda\alpha + (1-\lambda)\beta) \in \mathrm{epi} f.$$

Viceversa, se $\mathrm{epi} f$ è convesso, poiché

$$(x, f(x)), (y, f(y)) \in \mathrm{epi} f$$

si ha

$$(\lambda x + (1-\lambda)y, \lambda f(x) + (1-\lambda) f(y)) \in \mathrm{epi} f$$

e la 1.8. □

Osserviamo che il risultato precedente permette di riconoscere facilmente una funzione convessa.

È utile osservare che la 1.8 può essere scritta in diversi modi tutti utili per comprendere le proprietà delle funzioni convesse.

$$f(ty + (1-t)x) \leq tf(y) + (1-t)f(x) \tag{1.9}$$

$$f(z) \leq tf(y) + (1-t)f(x) \tag{1.10}$$

$$f(z) \leq f(x) + (z-x)\frac{f(y)-f(x)}{y-x} \tag{1.11}$$

Dalla definizione di convessità 1.10 si ricava sottraendo ad ambo i membri $f(y)$

$$f(z) - f(y) \leq (t-1)(f(y) - f(x)) \tag{1.12}$$

$$f(z) - f(y) \leq \frac{z-y}{y-x}(f(y) - f(x)) \tag{1.13}$$

$$\frac{f(z)-f(y)}{z-y} \geq \frac{f(y)-f(x)}{y-x} \tag{1.14}$$

una funzione f è convessa se e solo se

$$\frac{f(z)-f(y)}{z-y} \geq \frac{f(y)-f(x)}{y-x}$$

per ogni $x < z < y$

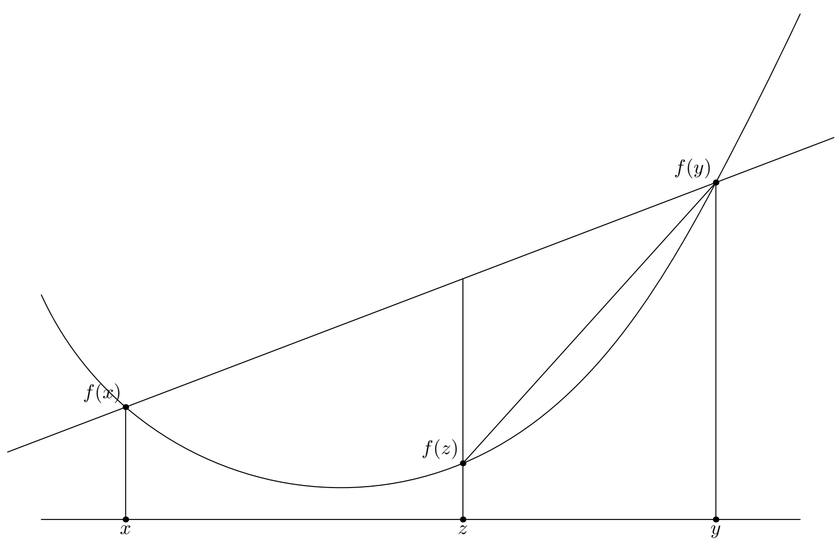


Figure 1.6:

Possiamo pertanto concludere, osservando che abbiamo sempre operato trasformando una disuguaglianza in una equivalente, che

Teorema 1.3 *f è convessa se e solo se*

$$r_x(h) = \frac{f(x+h) - f(x)}{h}$$

è crescente per ogni $x \in (a, b)$.

Dimostrazione. Basta osservare che la crescenza di r_x è equivalente alla 1.14, e quindi alla convessità. □

Teorema 1.4 *Se f è convessa allora, per ogni $x \in (a,b)$, esistono $f'_+(x)$ e $f'_-(x)$. Inoltre*

$$f'_-(x) \le f'_+(x).$$

Dimostrazione. La funzione r_x è crescente e, se δ è scelto in modo che

$$a < x - \delta < x < x + \delta < b$$

si ha

$$r_x(-\delta) < r_x(h) < r_x(\delta)$$

se $|h| < \delta$.

Pertanto r_x è limitata, sia superiormente che inferiormente, ed è lecito affermare che

$$\lim_{h \to 0^-} r_x(h) \quad e \quad \lim_{h \to 0^+} r_x(h)$$

esistono e sono finiti. □

Teorema 1.5 *Se f è convessa in (a,b), allora f non è derivabile in un insieme $N \subset (a,b)$ che risulta al più numerabile.*

Dimostrazione. Dal momento che $f'_-(x)$ e $f'_+(x)$ esistono finiti, per ogni $x \in (a,b)$ e si ha $f'_-(x) \le f'_+(x)$, possiamo affermare che

$$N = \{x \in (a,b) \ : \ f'_-(x) < f'_+(x)\ \}\ .$$

E' pertanto possibile associare ad ogni $x \in N$ un numero razionale q_x tale che

$$f'_-(x) < q_x < f'_+(x)$$

ed è immediato provare che, se $x \neq y$, si ha $q_x \neq q_y$.

Ciò ed il fatto che l'insieme dei numeri razionali è numerabile, permettono di concludere. □

D'altro canto, sempre dalla 1.10 f è convessa se e solo se:

$$f(z) \le tf(y) + (1-t)f(x) \tag{1.15}$$

$$f(z)(t + (1-t)) \le tf(y) + (1-t)f(x) \tag{1.16}$$

$$t(f(z) - f(y)) \le (1-t)(f(x) - f(z)) \tag{1.17}$$

$$(z-x)(f(z) - f(y)) \le (y-z)(f(x) - f(z)) \tag{1.18}$$

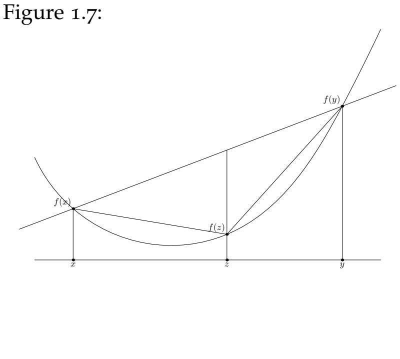
$$\frac{f(y) - f(z)}{y - z} \ge \frac{f(z) - f(x)}{z - x} \tag{1.19}$$

una funzione f è convessa se e solo se

$$\frac{f(y) - f(z)}{y - z} \ge \frac{f(z) - f(x)}{z - x}$$

per ogni $x < z < y$ Si veda la Figura 1

Figure 1.7:

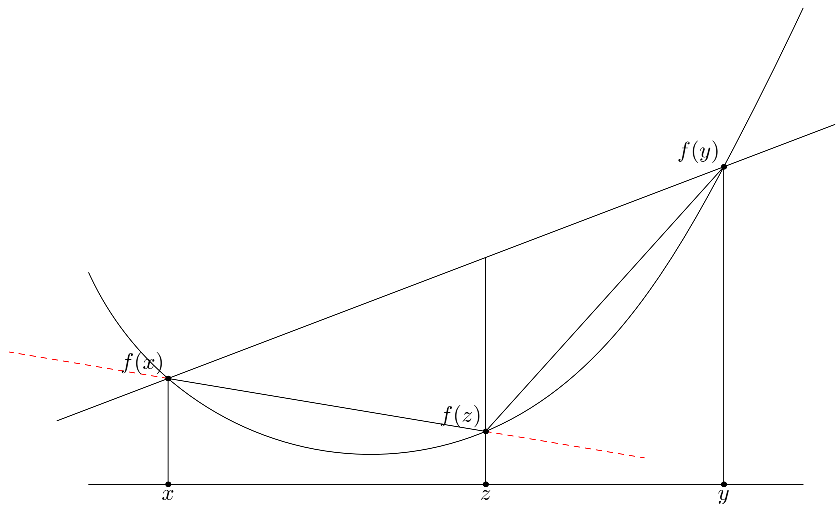


Per definizione f è convessa se il suo grafico sta, all'interno di un intervallo (x,y) contenuto nel suo dominio, sotto la retta secante il grafico che passa per $(x, f(x))$ ed $(y, f(y))$. La disuguaglianza 1.19 puo' essere riscritta come

$$\mathcal{F}(y) \ge f(z) + \frac{f(z) - f(x)}{z - x}(y - z)$$

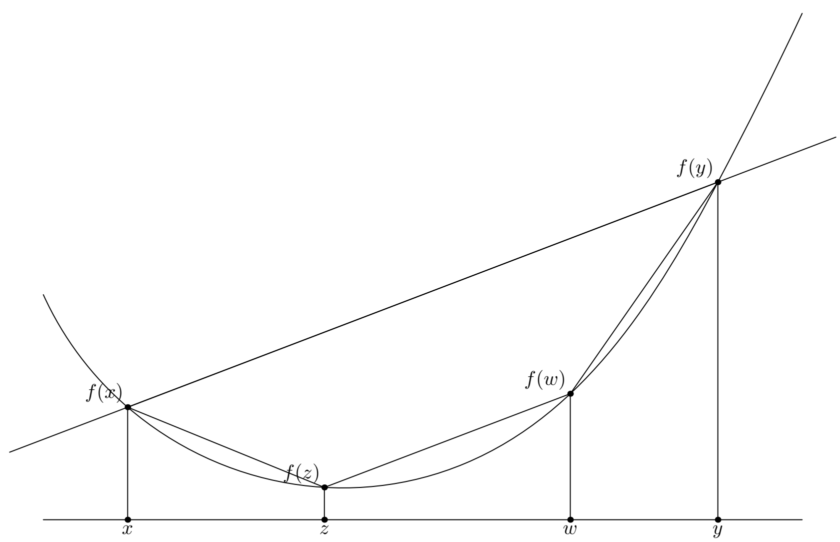
e mostra anche che il grafico di una funzione convessa sta, all'esterno di un intervallo (x, y) contenuto nel suo dominio, sopra la retta secante il grafico che passa per $(x, f(x))$ ed $(y, f(y))$.

Figure 1.8:



Ora, se $x < z < w < y$ si ha

Figure 1.9:



$$\frac{f(z) - f(x)}{z - x} \leq \frac{f(w) - f(z)}{w - z} \leq \frac{f(y) - f(w)}{y - w} \tag{1.20}$$

Passando al limite per $x \to z^-$ e per $y \to w^+$ se f è convessa e derivabile allora

$$f'(z) \leq f'(w) \tag{1.21}$$

e quindi f' è crescente.

Viceversa se f è derivabile ed f' è crescente, usando il teorema di Lagrange si può affermare che

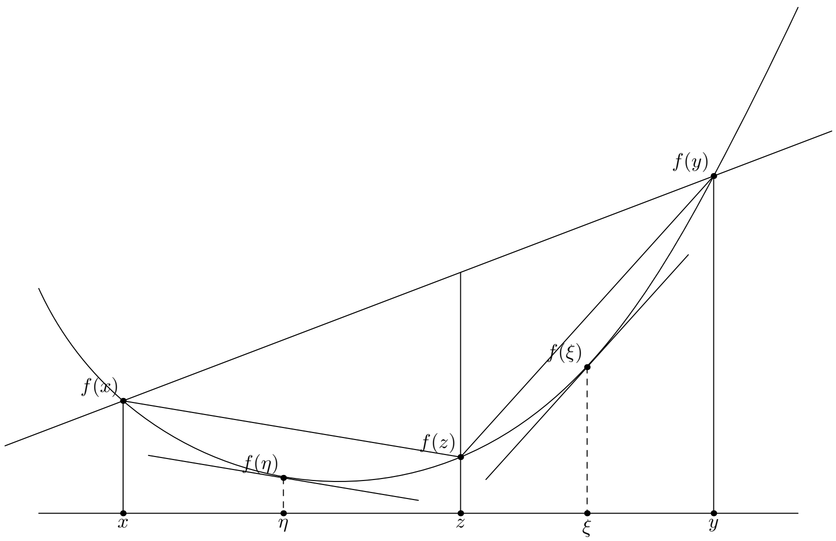


Figure 1.10:

$$\frac{f(z)-f(x)}{z-x} = f'(\xi) \geq f'(\eta) = \frac{f(y)-f(w)}{y-w} \tag{1.22}$$

e quindi f è convessa.

Ne concludiamo che:

Teorema 1.6 *Sia f derivabile, allora*

Sono fatti equivalenti

- *f è convessa in (a, b)*
- *f' è una funzione crescente in (a, b)*

Osserviamo infine che, se f è convessa, allora

$$f(y) - f(z) \geq (y-z)\frac{f(z)-f(x)}{z-x} \tag{1.23}$$

$$f(y) \geq f(z) + (y-z)\frac{f(z)-f(x)}{z-x} \tag{1.24}$$

e passando al limite per $x \to z$

$$f(y) \geq f(z) + f'(z)(y-z) \tag{1.25}$$

$$\tag{1.26}$$

e pertanto il grafico di f sta' sopra al grafico di ogni sua retta tangente, Se viceversa il grafico di f sta' sopra al grafico di ogni sua retta tangente, allora

$$f(y) \geq f(z) + f'(z)(y - z) \tag{1.27}$$

e

$$f(x) \geq f(z) + f'(z)(x - z) \tag{1.28}$$

da cui, tenendo conto che $y - z > 0$, e $x - z < 0$

$$\frac{f(y) - f(z)}{y - z} \geq f'(z) \geq \frac{f(z) - f(x)}{z - x} \tag{1.29}$$

e

$$f(y) - f(z) \geq (y - z)\frac{f(z) - f(x)}{z - x} \tag{1.30}$$

e quindi f è convessa.

Ne concludiamo che:

Teorema 1.7 *Sia f è derivabile, allora Sono fatti equivalenti*

- *f è convessa in (a, b)*
- *il grafico di f sta' sopra al grafico di ogni sua retta tangente*

I risultati che legano segno della derivata e crescenza della funzione permettono poi di concludere che

Teorema 1.8 *Sia f una funzione derivabile due volte in (a, b); sono condizioni equivalenti:*

- *f è convessa in (a, b);*
- *f' è crescente in (a, b);*
- *f'' è non negativa in (a, b).*

Definizione 1.4 *Sia $f : (a, b) \longrightarrow \mathbb{R}$, diciamo che f è concava in (a, b) se $-f$ è convessa in (a, b).*

Definizione 1.5 *Diciamo che $f : (a, b) \longrightarrow \mathbb{R}$ ha un punto di flesso in $x_0 \in (a, b)$ se esiste $\delta > 0$ tale che f è convessa (concava) in $(x_0 - \delta, x_0)$ e concava (convessa) in $(x_0, x_0 + \delta)$.*

Semplici esempi mostrano come sia possibile per una funzione avere un punto di flesso in 0 e

- non essere derivabile in 0 ($f(x) = \sqrt[3]{x}$)

- avere derivata non nulla in 0 ($f(x) = \sin x$)
- avere derivata nulla in 0 ($f(x) = x^3$).

Teorema 1.9 *Sia $f : (a, b) \longrightarrow \mathbb{R}$ e sia $x_0 \in (a, b)$, supponiamo f derivabile in (a, b); allora x_0 è un punto di flesso se e solo se f' è crescente (decrescente) in un intorno destro di x_0 e decrescente (crescente) in un intorno sinistro.*

E′ pertanto evidente che non è possibile caratterizzare un punto di flesso facendo uso soltanto della derivata prima nel punto.

Possiamo tuttavia provare i seguenti fatti

Teorema 1.10 *Sia $f : (a, b) \longrightarrow \mathbb{R}$ e sia $x_0 \in (a, b)$, supponiamo f derivabile in (a, b); allora x_0 è un punto di flesso se e solo se f' è crescente (decrescente) in un intorno destro di x_0 e decrescente (crescente) in un intorno sinistro.*

Corollario 1.1 *Sia $f : (a, b) \longrightarrow \mathbb{R}$ derivabile n volte in (a, b); allora x_0 è un punto di flesso per f solo se la prima derivata non nulla in x_0 è di ordine dispari.*

2. Estremi Relativi e Asintoti.

Abbiamo già visto cosa si intende per minimo e massimo assoluto di una funzione e abbiamo già trovato condizioni necessarie e sufficienti per l'esistenza di un minimo o un massimo assoluto. (Si veda il lemma **??** ed il teorema **??**).

In questo paragrafo ci occuperemo di stabilire la definizione di massimo e minimo relativo per una funzione e daremo condizioni necessarie e sufficienti per l'esistenza di un punto di minimo o di massimo relativo.

Definizione 2.1 *Sia $f : D \longrightarrow \mathbb{R}$ diciamo che $x_0 \in D$ è un punto di minimo (massimo) relativo per la funzione f se $\exists\delta > 0$ tale che se $x \in D \cap (x_0 - \delta, x_0 + \delta)$ si ha*

$$f(x) \geq f(x_0) \qquad\qquad (f(x) \leq f(x_0))$$

Proviamo ora condizioni necessarie e sufficienti per l'esistenza di un minimo o di un massimo relativo.

Essendo evidente l'analogia, ci occuperemo solo del caso dei minimi relativi.

Teorema 2.1 *Sia $f : (a,b) \longrightarrow \mathbb{R}$ una funzione derivabile n volte e sia $x_0 \in (a,b)$; sia k il primo numero naturale tale che $f^{(k)}(x_0) \neq 0$; allora x_0 è punto di minimo relativo per f se e solo se k è pari e $f^{(k)}(x_0) > 0$.*

Dimostrazione. Sia k il più piccolo numero naturale per cui $f^{(k)}(x_0) \neq 0$. Allora, usando la formula di Taylor con il resto di Peano si ha

$$f(x) = f(x_0) + \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + (x-x_0)^k\omega(x-x_0).$$

Pertanto

$$\frac{f(x)-f(x_0)}{(x-x_0)^k} = \frac{f^{(k)}(x_0)}{k!} + \omega(x-x_0).$$

Poichè

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{(x-x_0)^k} = \frac{f^{(k)}(x_0)}{k!}$$

per il teorema della permanenza del segno è lecito supporre che in un di x_0
$$\frac{f(x)-f(x_0)}{(x-x_0)^k}$$
abbia segno costante.

Ora, se x_0 è di minimo relativo deve essere k pari ed inoltre $f^{(k)}(x_0) > 0$; mentre se $f^{(k)}(x_0) > 0$ e k è pari, si vede, usando il teorema della permanenza del segno, che x_0 è un punto di minimo relativo. □

Più in generale è interessante osservare che, se P_n è il polinomio di Taylor di f centrato in x_0 di grado n, si ha

$$f(x) = P_n(x) + (x-x_0)^n \omega(x-x_0) \tag{2.1}$$

e se definiamo

$$P_n^1(x) = \sum_{k=1}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

si avrà

$$f(x) - f(x_0) = P_n^1(x) + (x-x_0)^n \omega(x-x_0) \tag{2.2}$$

mentre se

$$P_n^2(x) = \sum_{k=2}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

si avrà

$$f(x) - f(x_0) - f'(x_0)(x-x_0) = P_n^2(x) + (x-x_0)^n \omega(x-x_0) \tag{2.3}$$

Osserviamo che P^1 e P^2 sono, rispettivamente, i polinomi di Taylor di $f(x) - f(x_0)$ e $f(x) - f(x_0) - f'(x_0)(x-x_0)$.

Dividendo le 2.1,2.2,2.3 per P, P^1 e P^2, rispettivamente, otteniamo

$$\frac{f(x)}{P_n(x)} = 1 + \frac{(x-x_0)^n}{P_n(x)}\omega(x-x_0) \tag{2.4}$$
$$\frac{f(x)-f(x_0)}{P_n^1(x)} = 1 + \frac{(x-x_0)^n}{P_n^1(x)}\omega(x-x_0) \tag{2.5}$$
$$\frac{f(x)-f(x_0)-f'(x_0)(x-x_0)}{P_n^2(x)} = 1 + \frac{(x-x_0)^n}{P_n^2(x)}\omega(x-x_0) \tag{2.6}$$
$$\tag{2.7}$$

Poichè P_n,P_n^1,P_n^2, sono polinomi di grado n e quindi sono infinitesimi, per $x \to x_0$ di ordine al più n, tenendo conto che ω è a sua volta infinitesima, possiamo dedurre che

$$\frac{(x-x_0)^n}{P_n(x)}\omega(x-x_0) \quad \frac{(x-x_0)^n}{P_n^1(x)}\omega(x-x_0) \quad \frac{(x-x_0)^n}{P_n^2(x)}\omega(x-x_0) \tag{2.8}$$

sono infinitesimi per $x \to x_0$.

Il teorema della permanenza del segno permette quindi di affermare che

In un intorno di x_0

1. f ha lo stesso segno di P
2. $f(x) - f(x_0)$ ha lo stesso segno di P^1
3. $f(x) - f(x_0) - f'(x_0)(x - x_0)$ ha lo stesso segno di P^2

Poichè il segno di P_n in un intorno di x_0 è quello di $f(x_0)$, la prima affermazione si riduce semplicemente alla riaffermazione del teorema della permanenza del segno, tuttavia le altre due forniscono utili informazioni su crescenza e convessità.

Infatti poichè $f(x) - f(x_0)$ ha lo stesso segno di P_n^1 in un intorno di x_0 possiamo dire che x_0 è un punto di minimo relativo se siamo in grado di stabilire che P_n^1 è positivo in un intorno di x_0, viceversa possiamo dire che x_0 non è di minimo relativo se il polinomio P_n^1 cambia segno in un intorno di x_0.

Ora se supponiamo che f sia derivabile almeno n volte in $(a, b) \ni x_0$ e che $f^{(n)}(x_0)$ sia la prima derivata non nulla di f in x_0 possiamo considerare il polinomio P_n^1 che risulta essere definito da

$$P_n^1(x) = \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

e quindi risulta evidente che P_n^1 mantiene segno costante o cambia segno in un intorno di x_0 a seconda che n sia pari o dispari; nel caso che n sia pari il segno di P_n^1 è determinato dal segno di $f^{(n)}(x_0)$

In maniera simile $f(x) - f(x_0) - f'(x_0)(x - x_0)$ ha lo stesso segno di P_n^2 in un intorno di x_0 e quindi si ha che che x_0 è un punto di flesso se P_n^2 cambia segno in un intorno di x_0, viceversa possiamo dire che x_0 non è un punto di flesso se il polinomio P_n^2 è positivo in un intorno di x_0,

Ora se, come prima, supponiamo che f sia derivabile almeno n volte in $(a, b) \ni x_0$ e che $f^{(n)}(x_0)$ sia la prima derivata non nulla di f in x_0 possiamo considerare il polinomio P_n^2 che risulta essere definito da

$$P_n^2(x) = \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

e quindi risulta evidente che P_n^2 mantiene segno costante o cambia segno in un intorno di x_0 a seconda che n sia pari o dispari; nel caso che n sia pari il segno di P_n^2 è determinato dal segno di $f^{(n)}(x_0)$

Possiamo allora enunciare il seguente risultato

Teorema 2.2 *Sia $f : (a, b) \longrightarrow \mathbb{R}$ una funzione derivabile almeno n volte e sia $x_0 \in (a, b)$; sia $f^{(n)}(x_0) \neq 0$ la prima derivata che non si annulla,*

$n \geq 2$; allora x_0 è punto di flesso per f se e solo se n è dispari. Il segno di $f^{(n)}(x_0)$ fornisce poi informazioni sul fatto che il grafico di f sia sopra (funzione localmente convessa) o sotto (funzione localmente concava) la retta tangente al suo grafico

Definizione 2.2 *Siano $f, g : (a, +\infty) \longrightarrow \mathbb{R}$; diciamo che f e g sono asintotiche se*

$$\lim_{x \to +\infty} f(x) - g(x) = 0.$$

Nel caso in cui sia

$$g(x) = \alpha x + \beta$$

diciamo che g è un asintoto per f.

Teorema 2.3 *Sia $f : (a, +\infty) \longrightarrow \mathbb{R}$; la retta di equazione*

$$y = \alpha x + \beta$$

è un asintoto per f se e solo se

$$\alpha = \lim_{x \to +\infty} \frac{f(x)}{x} \qquad , \qquad \beta = \lim_{x \to +\infty} f(x) - \alpha x \tag{2.9}$$

Dimostrazione. E' immediato verificare che le 2.9 sono sufficienti affinché la retta sia asintoto.

Viceversa, se la retta è un asintoto, si ha □

$$\lim_{x \to +\infty} \frac{f(x) - \alpha x - \beta}{x} = \lim_{x \to +\infty} \frac{f(x)}{x} - \alpha = 0 \tag{2.10}$$

Definizione 2.3 *Sia $f : (a, b) \longrightarrow \mathbb{R}$, diciamo che la retta di equazione $x = c$ è un asintoto verticale per f se*

$$\lim_{x \to c} |f(x)| = +\infty$$

3. Ricerca Numerica di Zeri e Minimi.

Una delle applicazioni più tipiche della convessità consiste nella ricerca approssimata degli zeri di una funzione.

Il più semplice dei metodi di ricerca degli zeri è indubbiamente il metodo di bisezione di cui abbiamo già dato una dimostrazione in **??**

Il metodo di bisezione offre indubbi vantaggi di semplicità di applicazione e necessita di ipotesi ridotte alla sola continuità della funzione f; tuttavia, in presenza di migliori condizioni, si possono trovare metodi che convergono alla soluzione molto più velocemente.

Tali metodi, usualmente utilizzano la convessità della funzione, e sono tanto più importanti quanto è più grande la difficoltà di svolgere calcoli.

Chiaramente, con tempi di calcolo sempre più ridotti, tali metodi perdono parte della loro attrattiva anche se rimangono interessanti per la loro eleganza ed efficienza.

E' questo il caso del metodo di Newton (o delle tangenti) e del metodo della 'regula falsi'; essi convergono se le funzioni di cui si ricercano gli zeri sono convesse e possono essere generalizzati al caso non convesso purché le derivate prime e seconde della funzione f siano opportunamente maggiorabili o minorabili.

Cominciamo con l'occuparci dei metodi di ricerca degli zeri di una funzione.

Teorema 3.1 *- Metodo di Newton (o delle tangenti)- Supponiamo* $f : (\alpha, \beta) \longrightarrow \mathbb{R}$*, convessa e derivabile due volte in* (α, β)*; supponiamo inoltre che* $\alpha < a < b < \beta$ *e sia*

$$f(a) < 0 \;\; , \;\; f(b) > 0 \;\; .$$

Allora esiste uno ed un solo punto $c \in (a, b)$ *tale che*

$$f(c) = 0 \;\; , \;\; f'(x) \geq f'(c) > 0 \;\; \forall x \in [c, b].$$

Definiamo la successione x_n *nella seguente maniera:*

$$x_0 = b$$

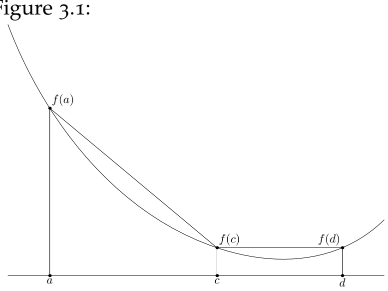
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad ;$$

allora:

- x_n *è decrescente e inferiormente limitata,*
- $\lim x_n = c$
- *se* $0 \leq f''(x) \leq M$ *per ogni* $x \in [a,b] \ni$ *se* $f'(a) = P > 0$ *si ha*

$$0 \leq x_n - c \leq \frac{2P}{M}\left(\frac{M}{2P}(b-a)\right)^{2^n}$$

Figure 3.1:



Dimostrazione. L'esistenza di c è immediata conseguenza del teorema degli zeri; per quel che riguarda l'unicità, se esistessero $c, d \in (a,b)$ tali che $c < d$, $f(c) = f(d) = 0$, per la 1.19 si avrebbe $f(a) \geq 0$.

Inoltre si ha $f'(c) > 0$; infatti se fosse $f'(c) \leq 0$, dal momento che f' è crescente in $[a,b]$, si avrebbe

$$f'(x) \leq f'(c) \leq 0 \quad \forall x \in [a,c] \quad ;$$

pertanto f sarebbe decrescente in $[a,c]$ e

$$f(a) \geq f(c) = 0 \,.$$

Pertanto

$$f'(x) \geq f'(c) > 0 \quad \forall x \in [c,b] \quad .$$

Proviamo ora per induzione che

$$c \leq x_{n+1} \leq x_n \leq b \quad .$$

Per ipotesi $c < x_0 = b$. Proviamo pertanto che, supponendo $c \leq x_n \leq b$ si può dedurre $c \leq x_{n+1} \leq x_n \leq b$. Sia

$$s(x) = f(x_n) + f'(x_n)(x - x_n)$$

allora

$$s(c) = f(x_n) + f'(x_n)(c - x_n) \leq f(c) = 0$$

e

$$s(x_n) = f(x_n) \geq 0 \quad .$$

Inoltre

$$s(x) = 0 \quad \Leftrightarrow \quad x = x_{n+1} \quad ;$$

ciò permette di concludere.

Pertanto la successione x_n è decrescente ed inferiormente limitata e si ha che $x_n \to \ell$, $\ell \in [a,b]$.

Per la continuità di f ed f' si ha $f'(\ell) > 0$,

$$\ell = \ell - \frac{f(\ell)}{f'(\ell)} \qquad \text{e} \qquad f(\ell) = 0.$$

Per l'unicità dello zero di f ne segue $c = \ell$.

Supponiamo infine vere le ipotesi del punto 3); usando la formula di Taylor con il resto di Lagrange si ottiene

$$0 = f(c) = f(x_n) + f'(x_n)(c - x_n) + \frac{f''(c_n)}{2}(c - x_n)^2 \quad , \quad c < c_n < x_n \quad ;$$

e, se teniamo conto che

$$f'(x_n)(x_n - x_{n+1}) = f(x_n)$$

otteniamo

$$0 = f'(x_n)(x_n - x_{n+1}) + f'(x_n)(c - x_n) + \frac{f''(c_n)}{2}(c - x_n)^2$$

e

$$x_{n+1} - c \;=\; \frac{f''(c_n)(c - x_n)^2}{2\, f'(x_n)}.$$

Pertanto

$$0 \leq x_{n+1} - c \leq \frac{M}{2P}(x_n - c)^2$$

e si può facilmente provare per induzione la tesi della 3). □

Il precedente metodo può essere generalizzato al caso in cui la funzione non sia convessa, ma siano verificate opportune condizioni.

Teorema 3.2 *- Metodo delle tangenti generalizzato - Supponiamo f derivabile due volte in $[2a - b, b]$ e sia $f(a) < 0$, $f(b) > 0$. Supponiamo inoltre che*

$$|f''(x)| \leq M \quad , \quad f'(x) \geq P > 0 \quad \forall x \in [2a - b, b]$$

e

$$\frac{M}{2P}(b - a) < 1 \quad .$$

Allora esiste un unico $c \in (a, b)$ tale che $f(c) = 0$; inoltre se definiamo x_n nella seguente maniera:

$$x_0 = b$$

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

si ha

1. $\lim\ x_n = c$
2. $|x_n - c| \leq \frac{2P}{M}\left(\frac{M}{2P}(b - a)\right)^{2^n}$.

Dimostrazione. Esistenza ed unicità di c sono ovvie poiché f è continua e strettamente crescente in $[a,b]$.

Si ha inoltre

$$0 = f(c) = f(x_n) + f'(x_n)(c - x_n) + \frac{f''(c_n)}{2}(c - x_n)^2,$$
$$|c_n - c| \leq |x_n - c|.$$

Si ottiene pertanto, come nel teorema precedente,

$$|x_{n+1} - c| \leq \frac{|f''(c_n)|}{2f'(x_n)}(x_n - c)^2 \quad .$$

Proviamo ora per induzione che

$$|x_n - c| \leq b - c \quad .$$

Innanzi tutto si ha

$$|x_0 - c| = b - c \quad ;$$

inoltre, se $|x_n - c| \leq b - c$ si può affermare che $2a - b \leq x_n \leq b$,

$$|c_n - c| \leq b - c \qquad \text{e} \qquad 2a - b \leq c_n \leq b$$

per cui

$$|x_{n+1} - c| \leq \frac{M}{2P}(x_n - c)^2 \leq \frac{M}{2P}(b - a)|x_n - c| \leq |x_n - c| \leq b - c.$$

Pertanto, dal momento che $x_n, c_n \in [2a - b, b]$ si ha che

$$|x_{n+1} - c| \leq \frac{M}{2P}(x_n - c)^2$$

e per induzione si deduce la tesi. □

Teorema 3.3 *-Metodo della regula falsi - Sia $f : (\alpha, \beta) \longrightarrow \mathbb{R}$ una funzione convessa e derivabile due volte in (α, β) e siano $a, b \in (\alpha, \beta)$, $a < b$, tali che $f(a) < 0$, $f(b) > 0$.*

Allora esiste uno ed un solo $c \in (a, b)$, tale che $f(c) = 0$ e $f'(x) \geq f'(c) > 0 \ \ \forall x \in [c, b]$.

Inoltre, se definiamo una successione x_n nella seguente maniera:

$$x_0, x_1 \in [c, b] \ , \ x_1 < x_0 \tag{3.1}$$

$$\begin{aligned} x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \\ = x_{n-1} - f(x_{n-1})\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \end{aligned} \tag{3.2}$$

si ha

- *x_n è decrescente ed inferiormente limitata,*
- $\lim\ x_n = c$
- *se $M, P \in \mathbb{R}$ sono tali che*

  $$0 \leq f''(x) \leq M \ \ , \ \ f'(x) \geq P > 0 \ \ \forall x \in [a, b]$$

  allora

  $$0 \leq x_n - c \leq \frac{2P}{M}\left(\frac{M}{2P}(b-a)\right)^{\delta_n}$$

  ove δ_n è la successione di Fibonacci.

Dimostrazione.

La prima parte dell'enunciato è comune al metodo delle tangenti di Newton e pertanto la dimostrazione è identica a quella del teorema 14.2.

Passiamo ora a provare la seconda parte; proviamo cioè, per induzione, che la successione x_n è decrescente ed inferiormente limitata da c.

Intanto si ha

$$c \leq x_1 \leq x_0 \leq b \ \ ;$$

dimostriamo perciò che, se $c \leq x_n \leq x_{n-1} \leq b$, si ha

$$c \leq x_{n+1} \leq x_n \leq b \ \ .$$

Definiamo

$$s(x) = f(x_n) + (x - x_n)\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$

allora (teorema 13.7)

$$s(c) = f(x_n) - (c - x_n)\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} \leq f(c) = 0$$

e

$$s(x_n) = f(x_n) > 0 \quad ,$$

inoltre

$$s(x) = 0 \quad \Leftrightarrow \quad x = x_{n+1} \ \ .$$

Se ne decuce che $c \leq x_{n+1} \leq x_n \leq x_{n-1} \leq b$.

Ora, essendo x_n decrescente e inferiormente limitata, posto

$$\ell = \lim x_n \in \mathbb{R}$$

si ha, per la continuità di f ed f' e per il fatto che

$$\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} = f'(c_n) \quad , \quad x_n \leq c_n \leq x_{n-1} \quad ,$$

che
$$\ell = \ell - \frac{f(\ell)}{f'(\ell)} \quad .$$

Ma $\ell \geq c$ in quanto $x_n \geq c$; pertanto $f'(\ell) \geq f'(c) > 0$, per cui si può asserire che $f(\ell) = 0$ e, per l'unicità dello zero,
$$\ell = c.$$

Supponiamo ora che valgano le ipotesi del punto 3); si ha

$$\begin{aligned} x_{n+1} - c = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} - c = \\ = (x_n - c)\left(1 - \frac{f(x_n) - f(c)}{x_n - c}\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}\right) = \\ = (x_n - c)(x_{n-1} - c)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}\frac{1}{x_{n-1} - c}\cdot \\ \cdot\left(\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} - \frac{f(x_n) - f(c)}{x_n - c}\right) \quad . \end{aligned}$$

Ora
$$\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} = \frac{1}{f'(c_n)} \quad , \quad x_n < c_n < x_{n-1}$$
mentre
$$\begin{aligned} \frac{1}{x_{n-1} - c}\left(\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} - \frac{f(x_n) - f(c)}{x_n - c}\right) = \\ = \frac{F(x_{n-1}) - F(c)}{x_{n-1} - c} \end{aligned}$$

dove
$$F(x) = \begin{cases} \frac{f(x) - f(x_n)}{x - x_n} & \text{se} \neq x_n \\ f'(x_n) & \text{se} x = x_n \end{cases}$$
inoltre
$$\frac{F(x_{n-1}) - F(c)}{x_{n-1} - c} = F'(\xi_n) \quad , \quad c < \xi_n < x_{n-1}$$
e
$$\begin{aligned} F'(\xi_n) = \frac{f'(\xi_n)(\xi_n - x_n) - f(\xi_n) + f(x_n)}{(\xi_n - x_n)^2} = -\frac{f''(d_n)}{2}, \\ c < d_n < x_n. \end{aligned}$$

Si ottiene quindi che
$$x_{n+1} - c = (x_n - c)(x_{n-1} - c)\frac{f''(d_n)}{2f'(c_n)} \leq \frac{M}{2P}(x_n - c)(x_{n-1} - c)$$

e se definiamo

$$\varepsilon_n = x_n - c$$

avremo

$$\varepsilon_0 = x_0 - c \geq x_1 - c = \varepsilon_1$$

ed anche

$$\varepsilon_{n+1} \leq \varepsilon_n \varepsilon_{n-1} \frac{M}{2P}$$

da cui si può dedurre per induzione che

$$\varepsilon_n \leq \left(\frac{M}{2P}\right)^{\delta_n - 1} \varepsilon_0^{\delta_n}$$

e la tesi. □

Teorema 3.4 *Metodo della regula falsi generalizzato - Sia f derivabile due volte in $[2a-b, b]$ e sia $f(a) < 0$, $f(b) > 0$. Supponiamo inoltre che*

$$|f''(x)| \leq M \quad , \quad f'(x) \geq P > 0 \qquad \forall x \in [2a-b, b]$$

e

$$\frac{M}{2P}(b-a) < 1 \quad .$$

Allora esiste uno ed un solo $c \in (a, b)$ tale che $f(c) = 0$. Inoltre, se definiamo x_n come nel teorema precedente, si ha

1. $\lim x_n = c$
2. $|x_n - c| \leq \frac{2P}{M}\left(\frac{M}{2P}(b-a)\right)^{\delta_n}$

   dove δ_n è la successione di Fibonacci.

Dimostrazione. L'esistenza e l'unicità di c è ovvia dalla continuità e dalla stretta crescenza di f.

Come nel teorema precedente si prova che

$$|x_{n+1} - c| = |x_n - c||x_{n-1} - c| \frac{f''(d_n)}{2f'(c_n)}$$

essendo $|c_n - x_n| \leq |x_n - x_{n-1}|$, $|d_n - c| \leq |x_n - c|$.

Si può poi provare per induzione che

$$|x_n - c| \leq b - c$$

e ne segue

$$|x_{n+1} - c| \leq |x_n - c||x_{n-1} - c| \frac{M}{2P} \ .$$

Si conclude come nel caso convesso. □

Possiamo infine ricavare un metodo per individuare gli zeri di una funzione osservando che

$$f(x) = 0 \quad \Leftrightarrow \quad x = x + \alpha f(x) \quad .$$

Posto

$$F(x) = x + \alpha f(x)$$

il nuovo problema è risolvere l'equazione

$$F(x) = x$$

e viene usualmente indicato come ricerca di un *punto fisso* per la funzione F.

A questo proposito è immediato provare il seguente risultato:

Teorema 3.5 *Sia* $f : [a, b] \longrightarrow [a, b]$ *una funzione continua, allora esiste* $c \in [a, b]$ *tale che* $f(c) = c$ *.*

Dimostrazione. Definiamo $\phi : [a, b] \longrightarrow \mathbb{R}$ mediante la

$$\phi(x) = f(x) - x \quad .$$

Si ha

$$\phi(a) = f(a) - a \geq 0 \quad , \quad \phi(b) = f(b) - b \leq 0 \quad ;$$

pertanto applicando il teorema degli zeri, si ottiene

$$\exists c \in [a, b] \; : \; \phi(c) = f(c) - c = 0.$$

□

Possiamo anche provare il seguente risultato che permette di ricavare un semplicissimo algoritmo per approssimare la soluzione del problema considerato.

Teorema 3.6 - *Metodo delle contrazioni - Sia* $D = [x_0 - \rho, x_0 + \rho]$ *e supponiamo che* $f : D \longrightarrow \mathbb{R}$ *soddisfi le seguenti condizioni*

1. $|f(x) - f(y)| \leq k|x - y| \;\; , \forall x, y \in D \; , \; 0 \leq k < 1$.
2. $|x_0 - f(x_0)| \leq (1 - k)\rho$.

Allora esiste uno ed un solo punto $c \in D$ *tale che* $f(c) = c$.

Inoltre, definita una successione x_n*, mediante la*

$$x_{n+1} = f(x_n)$$

si ha

$$\lim \; x_n = c$$

e

$$|x_n - c| \leq \rho k^n.$$

Dimostrazione. Cominciamo con il provare che c, se esiste, è unico; se infatti esistesse $d \in D$ tale che $f(d) = d$ si avrebbe

$$|d - c| = |f(d) - f(c)| \leq k|d - c|$$

da cui

$$(1-k)|d-c| \leq 0$$

e dal momento che $1-k>0$ si ha $d=c$.

Vediamo ora che è possibile considerare la successione x_n; allo scopo sarà sufficiente provare per induzione che $x_n \in D \quad \forall n \geq 0$.

Osserviamo intanto che $x_0, x_1 \in D$ e dimostriamo che, supposto $x_k \in D$, $0 \leq k \leq n$, si ha $x_{n+1} \in D$.

Si ha

$$|x_{n+1} - x_n| = |f(x_n) - f(x_{n-1})| \leq k|x_n - x_{n-1}| \leq \ldots \leq k^n |x_1 - x_0|$$

e

$$\begin{aligned} |x_{n+1} - x_0| &\leq \sum_{i=0}^{n} |x_{i+1} - x_i| \leq \sum_{i=0}^{n} k^i |x_1 - x_0| = \\ &= \frac{1-k^{n+1}}{1-k} |x_1 - x_0| \leq (1-k^{n+1})\rho \leq \rho. \end{aligned}$$

Si ha inoltre

$$\begin{aligned} |x_{n+p} - x_n| &\leq \sum_{i=0}^{p-1} |x_{n+i+1} - x_{n+i}| \leq \sum_{i=0}^{p-1} k^{n+i} |x_1 - x_0| = \\ &= |x_1 - x_0| k^n \sum_{i=0}^{p-1} k^i = |x_1 - x_0| k^n \frac{1-k^p}{1-k} \leq \rho k^n. \end{aligned}$$

Ciò prova che x_n è una successione di Cauchy e pertanto esiste $c \in D$ tale che

$$c = \lim\ x_n \quad .$$

Per la continuità di f

$$c = \lim\ x_{n+1} = \lim\ f(x_n) = f(c) \quad .$$

Infine, passando al limite per $p \to +\infty$ nella precedente disuguaglianza si ha la tesi. □

Osservazione. Una funzione soddisfacente l'ipotesi 1) del teorema 14.7 si chiama *contrazione* di costante k.

Una funzione tale che

$$|f(x) - f(y)| \leq k|x-y| \quad , \quad k \geq 0$$

si chiama invece *lipschitziana* di costante k .

Osserviamo che una funzione lipschitziana è continua, ed inoltre f è lipschitziana se e solo se il suo rapporto incrementale è limitato. □

Possiamo più in generale provare che

Teorema 3.7 *Sia f una funzione continua e supponiamo che esista $m \in N$ tale che*

$$\phi = f^m$$

soddisfi le ipotesi del teorema precedente.

Allora esiste uno ed un solo punto $c \in D$ tale che $f(c) = c$.

Dimostrazione. Applicando il risultato precedente a ϕ si ottiene che esiste un unico punto $c \in D$ tale che $f^m(c) = \phi(c) = c$.

Inoltre si ha

$$\phi^p(f(c)) = f(\phi^p(c))$$

e quindi, dal momento che,

$$\phi^p(f(c)) \to c \quad \text{per} \;\; p \to +\infty$$

si ha

$$f(c) = f(\phi^p(c)) = \phi^p(f(c)) \to c$$

da cui

$$f(c) = c.$$

Pertanto ogni punto fisso di ϕ è anche punto fisso di f; il viceversa è di immediata verifica. □

Un semplice esempio di applicazione del precedente teorema è dato dalla funzione

$$f(x) = \text{atn}(x + \pi).$$

Infatti f non è contrazione, ma f^2 lo è.

Si osservi che un metodo rapido e semplice per verificare le ipotesi del teorema 14.7 è usare il teorema di Lagrange, con opportune maggiorazioni della derivata prima.

Illustriamo infine brevemente un metodo di ricerca degli zeri che si dimostra particolarmente efficace nell'uso pratico.

Sia $f : [a, b] \longrightarrow \mathbb{R}$ una funzione continua e supponiamo che f(a)$f(b) <$ 0; in modo che f ammetta almeno uno zero in (a, b).

Definiamo una successione x_n di punti di $[a, b]$ nella seguente maniera:

$$x_0 = a \;\;,\;\; x_1 = b \;\;,\;\; f_0 = f(x_0) \;\;,\;\; f_1 = f(x_1)$$

$$x_{n+1} = x_n - f_n \frac{x_n - x_{n-1}}{f_n - f_{n-1}} \;\;,\;\; f_{n+1} = f(x_{n+1})$$

inoltre se

$$f_n f_{n+1} > 0$$

si sostituisce

$$x_n \quad \text{con} \quad x_{n-1} \quad \text{e} \quad f_n \quad \text{con} \quad f_{n-1}/2.$$

Il metodo descritto permette di trovare un intervallo di ampiezza arbitrariamente piccola, di estremi x_{n-1} e x_n, entro il quale è certamente localizzato uno zero della funzione in esame; si può inoltre verificare una notevole efficacia, in termini di rapidità di convergenza del metodo stesso.

Molto spesso non è facile trovare i massimi e minimi relativi od assoluti di una funzione usando i metodi classici in quanto, ad esempio, non si possono facilmente valutare gli zeri della derivata prima o addirittura la funzione in esame non è derivabile. In tali casi si ricorre a metodi per la determinazione numerica approssimata dei punti e dei valori di minimo.

Tali metodi si prefiggono di restringere, fino alla lunghezza desiderata, l'intervallo in cui c'è il minimo della funzione: in altre parole, assegnata una funzione f su un intervallo [a,b], che ammetta ivi un minimo relativo interno, si cerca un intervallo più piccolo $[a_1, b_1] \subset [a, b]$ in modo che f ammetta ancora in $[a_1, b_1]$ un minimo relativo interno. Iterando il procedimento si può ridurre l'intervallo fino a trovare $[a_k, b_k]$ in modo che la sua ampiezza sia più piccola della precisione desiderata.

La possibilità di trovare l'intervallo $[a_1, b_1]$ a partire da $[a, b]$ è garantita da alcuni semplici risultati che illustriamo di seguito e dipende dalla scelta di opportuni punti in [a,b].

Lemma 3.1 *Sia $f : [a, b] \longrightarrow \mathbb{R}$ continua e supponiamo che esista $c \in (a, b)$ tale che*

$$f(c) \leq \min\{f(a), f(b)\} \quad .$$

Allora esiste un punto $x_0 \in (a, b)$ di minimo relativo per f.

Dimostrazione. Per il teorema di Weierstrass esiste $x_0 \in [a, b]$ tale che

$$f(x_0) \leq f(x) \quad \forall x \in [a, b]$$

e poiché

$$f(c) \leq \min\{f(a), f(b)\}$$

il minimo assoluto non può essere assunto soltanto in a o in b. □

Stabiliamo ora un criterio di scelta per passare dall'intervallo $[a, b] = [a_0, b_0]$ all'intervallo $[a_k, b_k]$ iterativamente.

Definizione 3.1 *Sia $f : [a, b] \longrightarrow \mathbb{R}$ continua, diciamo che è data una successione di intervalli $[a_k, b_k]$ localizzante un minimo relativo se:*

- *$[a_0, b_0] = [a, b]$;*
- *è assegnata una regola che permette di scegliere in (a_k, b_k) due punti, che indichiamo con α_k, β_k, (indichiamo questa regola con il nome di regola di scelta);*

- *è individuato* $[a_{k+1}, b_{k+1}]$ *nella seguente maniera*

$$[a_{k+1}, b_{k+1}] = [a_k, \beta_k] \qquad \text{se} \qquad f(\alpha_k) \leq f(\beta_k)$$
$$[a_{k+1}, b_{k+1}] = [\alpha_k, b_k] \qquad \text{se} \qquad f(\alpha_k) > f(\beta_k)$$

Diciamo inoltre che la regola di scelta è aurea se accade che

$$\beta_{k+1} = \alpha_k \quad \text{se} \;\; f(\alpha_k) \leq f(\beta_k)$$
$$\alpha_{k+1} = \beta_k \quad \text{se} \;\; f(\alpha_k) > f(\beta_k).$$

Osserviamo che una regola di scelta aurea permette di introdurre ad ogni iterazione un solo nuovo punto e quindi causa ad ogni passo la necessità di un solo nuovo calcolo di funzione anziché due.

Teorema 3.8 *Sia* $f : [a, b] \longrightarrow \mathbb{R}$ *continua e sia* $[a_k, b_k]$ *una successione di intervalli localizzante relativa ad una scelta di punti aurea; allora se*

$$\min\{f(\alpha_0), f(\beta_0)\} \leq \min\{f(a_0), f(b_0)\},$$

in ogni intervallo (a_k, b_k) *esiste un punto di minimo relativo per* f.

Dimostrazione. Se

$$f(\alpha_0) \leq f(\beta_0) \quad \text{si ha} \quad [a_1, b_1] = [a_0, \beta_0]$$

e

$$f(\alpha_0) \leq \min\{f(a_1), f(b_1)\} \;\; ;$$

se

$$f(\alpha_0) > f(\beta_0) \quad \text{si ha} \quad [a_1, b_1] = [\alpha_0, b_0]$$

e

$$f(\beta_0) \leq \min\{f(a_1), f(b_1)\} \;\; .$$

Poiché la regola di scelta è aurea risulta

$$\min\{f(\alpha_1), f(\beta_1)\} \leq \min\{f(a_1), f(b_1)\}$$

e iterando il procedimento, per il lemma 14.9, si ha la tesi. □

Corollario 3.1 *Se oltre alle ipotesi del teorema 14.11,* f *ammette un unico punto di minimo assoluto* $x_0 \in (a, b)$, *allora*

$$a_k \leq x_0 \leq b_k \quad \forall k \in N.$$

I risultati precedenti sottolineano l'importanza di conoscere una scelta di punti aurea, valutare l'ampiezza di $[a_k, b_k]$ e conoscere condizioni che garantiscano l'unicità del minimo di f.

Una buona scelta aurea dei punti α_k e β_k può essere fatta seguendo diversi criteri: innanzi tutto deve essere

$$\alpha_k < \beta_k$$

e

$$b_k - \alpha_k = \beta_k - a_k$$

in modo che i casi da considerare di volta in volta siano ridotti in numero e affinché l'ampiezza di ogni intervallo sia indipendente da ciò che accade nel precedente intervallo.

Ciò premesso imporremo

$$\alpha_{k+1} = \beta_k \quad \text{se } [a_{k+1}, b_{k+1}] = [\alpha_k, b_k]$$

e

$$\beta_{k+1} = \alpha_k \quad \text{se } [a_{k+1}, b_{k+1}] = [a_k, \beta_k] \quad .$$

Per scegliere α_k e β_k sotto queste condizioni sarà perciò opportuno scegliere θ_k in modo che

$$\alpha_k = a_k + (1 - \theta_k)(b_k - a_k)$$

$$\beta_k = b_k - (1 - \theta_k)(b_k - a_k)$$

ed imporre condizioni su θ_k .

Affinché $\alpha_k < \beta_k$ si deve avere

$$-1 + (1 - \theta_k) + (1 - \theta_k) < 0$$

e

$$1 - 2\theta_k < 0$$

da cui

$$\theta_k > \frac{1}{2}. \tag{14.1}$$

Affinché $\alpha_{k+1} = \beta_k$ quando $[a_{k+1}, b_{k+1}] = [\alpha_k, b_k]$ deve aversi

$$a_{k+1} + (1 - \theta_{k+1})(b_{k+1} - a_{k+1}) = b_k - (1 - \theta_k)(b_k - a_k)$$

$$\begin{aligned} a_k + (1 - \theta_k)(b_k - a_k) + (1 - \theta_{k+1})[b_k - a_k - (1 - \theta_k)(b_k - a_k)] = \\ = b_k - (1 - \theta_k)(b_k - a_k) \end{aligned}$$

da cui

$$\theta_{k+1} = \frac{1 - \theta_k}{\theta_k} \quad . \tag{14.2}$$

La stessa relazione si ottiene per avere $\alpha_k = \beta_{k+1}$ quando

$$[a_{k+1}, b_{k+1}] = [a_k, \beta_k].$$

La successione θ_k dovrà pertanto soddisfare le condizioni (14.1) e (14.2) e dipenderà dalla scelta del primo termine θ_1 (dal quale è univocamente determinata).

Ad ogni iterazione l'intervallo riduce la sua lunghezza di un fattore θ_k e dopo n iterazioni la lunghezza sarà ridotta del fattore

$$\prod_{k=1}^{n} \theta_k \quad .$$

Osserviamo che, se

$$\theta_k = \frac{\phi_k}{\psi_k}$$

deve aversi

$$\frac{\phi_{k+1}}{\psi_{k+1}} = \frac{\psi_k - \phi_k}{\phi_k}$$

e ciò è verificato se

$$\phi_{k+1} = \psi_{k-1} - \phi_k \quad e \quad \psi_{k+1} = \phi_k \quad , \quad \phi_0 = \psi_1 \; . \tag{14.3}$$

In tal caso si ha

$$\theta_k = \frac{\phi_k}{\phi_{k-1}}$$

e

$$\prod_{k=1}^{n} \theta_k \; = \frac{\phi_n}{\phi_0}.$$

Pertanto, dopo n iterazioni, la diminuzione della lunghezza dell'intervallo è di

$$\frac{\phi_n}{\phi_0} \quad ,$$

tale quantità dipendendo dai valori ϕ_0 e ϕ_1 che fino ad ora sono arbitrari.

Essendo

$$\frac{\phi_k}{\phi_0} = \frac{1}{2\sqrt{5}} \left((\sqrt{5}+1) \left(\frac{\sqrt{5}-1}{2} \right)^k + (\sqrt{5}-1) \left(-\frac{\sqrt{5}+1}{2} \right)^k \right) + \\ + \frac{\phi_1}{\phi_0} \frac{1}{\sqrt{5}} \left(\left(\frac{\sqrt{5}-1}{2} \right)^k - (-1)^k \left(\frac{\sqrt{5}+1}{2} \right)^k \right) ,$$

affinché ϕ_n / ϕ_0 sia minimo occorre che $\begin{array}{ll} \frac{\phi_1}{\phi_0} & \text{sia massimo, se } n \text{ è pari;} \\ \frac{\phi_1}{\phi_0} & \text{sia minimo, se } n \text{ è dispari.} \end{array}$

D'altra parte, affinché

$$\frac{1}{2} \leq \theta_n \leq 1$$

deve essere

$$\frac{1}{2} \leq \theta_{n-1} \leq \frac{2}{3}$$

$$\frac{3}{5} \leq \theta_{n-2} \leq \frac{2}{3}$$

e così di seguito.

Posto

$$r_k = \frac{F_k}{F_{k+1}}$$

dove F_k è la successione di Fibonacci, risulta $\begin{array}{ll} r_k \leq \theta_{n-k} \leq r_{k+1} & \text{se } k \text{ è dispari} \\ r_{k+1} \leq \theta_{n-k} \leq r_k & \text{se } k \text{ è pari} \end{array}$

e quindi

$r_{n-1} \leq \theta_1 \leq r_n$ se n è pari

$r_n \leq \theta_1 \leq r_{n-1}$ se n è dispari.

Tenendo conto che

$$\theta_1 = \frac{\phi_1}{\phi_0}$$

la massima riduzione dell'intervallo si ha per

$$\frac{\phi_1}{\phi_0} = r_n = \frac{F_n}{F_{n+1}}$$

per cui si può scegliere

$$\phi_1 = F_n \quad , \quad \phi_0 = F_{n+1}$$

e

$$\phi_k = F_{n-k}$$

non appena si ricordi la (14.3).

Con questa scelta

$$\prod_{k=1}^{n} \theta_k = \frac{1}{F_{n+1}} \quad .$$

Una scelta più banale per la successione θ_k è

$$\theta_k = \text{costante} = \theta$$

ed in tal caso affinché la (14.2) sia verificata deve risultare

$$\theta = \frac{\sqrt{5}-1}{2} \quad ;$$

dopo n iterazioni la riduzione dell'intervallo di partenza è proporzionale ad un fattore

$$\left(\frac{\sqrt{5}-1}{2}\right)^n .$$

Osserviamo che

$$\lim_n \; F_{n+1} \left(\frac{\sqrt{5}-1}{2}\right)^n = \frac{3+\sqrt{5}}{2\sqrt{5}} \approx 1.1708203932499369089227 5$$

per cui il metodo con θ costante, essendo di poco peggiore, ma estremamente più semplice, è consigliabile.

Va inoltre notato che se θ è costante, si può migliorare l'approssimazione a piacere, proseguendo nelle iterazioni; nell'altro caso invece il numero di iterazioni va fissato a priori, e per cambiarlo occorre rifare tutti calcoli dall'inizio.

Ai metodi qui illustrati vanno aggiunti quelli che cercano gli zeri della derivata, o del rapporto incrementale (si vedano i metodi di ricerca numerica degli zeri).

Notiamo fra l'altro che, se $f(x+h) - f(x) = 0$ allora $\exists c \in (x, x+h)$ punto di minimo o di massimo relativo per f.

Osserviamo esplicitamente che, a causa degli arrotondamenti con cui le macchine eseguono i calcoli, può accadere che una funzione venga valutata zero anche in punti 'non vicini' alla soluzione effettiva: ad esempio, utilizzando meno di 20 cifre significative $1 - \cos(x^{10})$ risulterà nulla per $x = .1$.