4
Bestiario

In questo capitolo presenteremo le principali classi di funzioni reali di variabile reale che si incontrano nella pratica scientifica. Vedremo anche alcune delle tecniche principali per trovare funzioni di un certo tipo che approssimano meglio possibile dei dati sperimentali.

4.1 Funzioni lineari

Le funzioni reali di variabile reale più semplici (dopo le costanti. . .) sono le funzioni lineari. Rappresentano relazioni di proporzionalità: una funzione è lineare se il suo valore varia in modo proporzionale alla variazione dell'argomento. In altre parole, una funzione f è lineare se esiste un numero reale $m \in \mathbb{R}$ (di solito non nullo) tale che se la variabile indipendente x varia di una quantità p allora la variabile dipendente $f(x)$ varia di mp.

Vediamo come si deduce da questa definizione la formula che descrive una funzione lineare. Supponiamo di variare il valore della variabile indipendente da x_0 a x; la variazione è quindi uguale a $p = x - x_0$, e tradizionalmente si indica con $\Delta x = x - x_0$. Se $f\colon \mathbb{R} \to \mathbb{R}$ è una funzione lineare, la variazione $\Delta f = f(x) - f(x_0)$ del suo valore deve soddisfare la relazione

$$\Delta f = m\,\Delta x\ .$$

Inserendo in questa formula le definizioni di Δf e Δx otteniamo

$$f(x) - f(x_0) = \Delta f = m\,\Delta x = m(x - x_0)\ ,$$

per cui

$$f(x) = mx + d\ , \tag{4.1}$$

con $d = f(x_0) - mx_0$.

Viceversa, supponiamo che la funzione $f\colon \mathbb{R} \to \mathbb{R}$ sia data dalla formula (4.1). Allora

$$\Delta f = f(x) - f(x_0) = mx + d - (mx_0 + d) = m(x - x_0) = m\Delta x\ ,$$

cioè f rappresenta una relazione di proporzionalità.

Riassumendo, le *funzioni lineari* (a volte chiamate anche *funzioni lineari affini,* riservando il termine "lineare" alle funzioni di questo tipo con $d = 0$) sono tutte e sole le funzioni della forma (4.1) per opportuni m, $d \in \mathbb{R}$.

Nella pratica sperimentale, capita spesso di trovare dati che dipendono in maniera lineare da una variabile (almeno per certi intervalli della variabile; vedi l'Osservazione 4.1); si pone quindi il problema di come trovare la legge che esprime questa relazione a partire dai dati sperimentali. In altre parole, vogliamo recuperare i coefficienti m e d conoscendo alcuni punti del grafico della funzione. Vediamo un esempio molto semplice, ma già significativo, di questa situazione.

Esempio 4.1 *È noto che la percentuale di semi che germogliano di una certa pianta dipende dalla temperatura. Per una determinata varietà di pomodoro, è stato verificato che alla temperatura di* 12 °C *germoglia il* 40% *dei semi, mentre alla temperatura di* 15 °C *germoglia il* 70% *dei semi. Trova la relazione fra la temperatura e la percentuale di semi germogliati, supponendo che si tratti di una relazione lineare.* Indichiamo con $P(T)$ la percentuale di semi che germoglia alla temperatura di T °C. Siccome abbiamo supposto che la funzione $P\colon \mathbb{R} \to \mathbb{R}$ che associa alla temperatura T la percentuale $P(T)$ sia lineare, possiamo scrivere

$$P(T) = mT + d$$

per opportuni m, $d \in \mathbb{R}$; il nostro obiettivo è usare i dati sperimentali per calcolare m e d. Noi sappiamo che $P(12) = 40$ e $P(15) = 70$; dunque

$$30 = 70 - 40 = P(15) - P(12) = \Delta P = m\Delta T = m(15 - 12) = 3m\ ,$$

(attenzione: in questo esempio la variabile indipendente si chiama T e la variabiile dipendente P, per cui abbiamo scritto ΔT e ΔP invece di Δx e Δf) da cui deduciamo

$$m = \frac{30}{3} = 10\ .$$

Per ricavare d basta notare che si deve avere

$$40 = P(12) = m \cdot 12 + d = 10 \cdot 12 + d = 120 + d\ ,$$

per cui

$$d = -80\ .$$

Quindi l'unica funzione lineare che rappresenta correttamente i dati sperimentali è

$$P(T) = 10\,T - 80\ . \tag{4.2}$$

Osservazione 4.1 È importante notare che per arrivare a questa soluzione abbiamo *supposto a priori* che la funzione da trovare fosse di tipo lineare; è un'ipotesi, e non una conseguenza. Del resto, da due sole coppie di dati è ben difficile immaginare, senza altre informazioni, quale possa essere l'andamento della funzione che volevamo studiare. Quale sia l'ipotesi giusta da fare (se lineare, quadratica, esponenziale o di altro tipo) può venire suggerito da quanto sappiamo sul fenomeno biologico che stiamo studiando. Altrimenti, conviene fare molte più misure e cercare di capire quale sia la funzione che *meglio approssima* i dati che abbiamo ottenuti, tenendo presente che le misure sono sicuramente affette da errori sperimentali. Ne parleremo in dettaglio nella Sezione 4.3.

Come già accennato nel capitolo precedente, lo scopo di ottenere una formula come la (4.2) è effettuare predizioni. Ci permette di dare risposte plausibili[1], senza bisogno di altre misure, a domande del tipo: quale percentuale di semi germoglierà alla temperatura di 14 °C? A quale temperatura germoglierà il 50% dei semi?

Esempio 4.2 Supponiamo quindi che la relazione fra la percentuale di semi che germogliano e la temperatura per questa varietà di pomodori sia data dalla formula (4.2). Allora la percentuale di semi che germogliano a 14 °C è

$$P(14) = 10 \cdot 14 - 80 = 140 - 80 = 60\% \,.$$

Trovare la temperatura T a cui germoglia il 50% dei semi equivale invece a risolvere l'equazione $P(T) = 50$, cioè

$$50 = P(T) = 10\,T - 80 \;;$$

quindi $10\,T = 130$, cioè $T = 13$ °C.

Esempio 4.3 Per la stessa varietà di pomodori, vogliamo trovare quale percentuale di semi germoglierà alla temperatura di 10 °C, e a quale temperatura germoglierà il 90% dei semi. La risposta alla prima domanda è $P(10) = 10 \cdot 10 - 80 = 20\%$, mentre per rispondere alla seconda domanda risolviamo l'equazione $P(T) = 90$ ottenendo $T = 17$ °C.

Le predizioni dell'Esempio 4.2 sono frutto di una *interpolazione*. Infatti, abbiamo dati sperimentali sia per valori della variabile indipendente inferiori a quelli coinvolti in queste predizioni, sia per valori superiori: sappiamo cosa succede a 12 e 15 °C, e deduciamo cosa accade a 13 e 14 °C. Invece, le predizioni dell'Esempio 4.3 sono frutto di una *estrapolazione:* i valori della variabile indipendente coinvolti nelle predizioni (10 e 17 °C) sono esterni all'intervallo dei valori della variabile indipendente per cui abbiamo dati sperimentali. Le estrapolazioni sono sempre molto più rischiose delle interpolazioni, in quanto l'ipotesi iniziale (che la relazione fosse di tipo lineare) potrebbe valere *solo all'interno di un determinato intervallo di valori.*

Esempio 4.4 Usando la (4.2) "prediciamo" che alla temperatura di 19 °C germoglierà il $P(19) = 10 \cdot 19 - 80 = 110\%$ dei semi, cosa piuttosto improbabile a meno

[1] Plausibili, e non certe: vedi l'Osservazione 4.2.

di generazione spontanea di nuovi semi dal nulla... Analogamente, la “predizione” che alla temperatura di 5 °C germogli il $P(5) = 10 \cdot 5 - 80 = -30\%$ dei semi ha ben poco senso.

Dunque quando si ipotizza un certo andamento per dei dati sperimentali è importante indicare con chiarezza l’intervallo dei valori per cui si ritiene valida l’ipotesi; al di fuori di quei valori l’estrapolazione potrebbe non avere senso anche se la funzione che rappresenta l’andamento dei dati è ancora definita.

Osservazione 4.2 Determinare l’intervallo dei valori in cui la formula ottenuta può essere valida è spesso un esercizio di buon senso: nel nostro caso, una percentuale maggiore del 100% o negativa non ha senso, per cui dobbiamo escludere i valori che danno risultati del genere. Nella pratica sperimentale, c’è però un ulteriore passo importante da fare: confrontare le predizioni sensate (le interpolazioni) ottenute con nuovi risultati sperimentali. Infatti, le nostre predizioni sono basate su un’ipotesi (che la funzione fosse di tipo lineare), ipotesi che dobbiamo verificare nei fatti. Se le nostre predizioni sono in buon accordo con le nuove misure (tenendo presente gli inevitabili errori sperimentali) allora possiamo dirci soddisfatti della nostra ipotesi; se invece non lo sono, dobbiamo cambiare ipotesi (vedi gli Esempio 4.7 e 4.9, e l’Esercizio 4.2 della Sezione 4.4).

Lasciamo ora crescere in pace i nostri pomodori, e vediamo come si affrontano in generale i problemi che abbiamo risolto in questo caso particolare. Supponiamo di avere due coppie $P_0 = (x_0, y_0)$ e $P_1 = (x_1, y_1)$ di dati; vogliamo trovare una funzione lineare $f(x) = mx + d$ tale che P_0 e P_1 appartengano al grafico di f, cioè tale che $f(x_0) = y_0$ e $f(x_1) = y_1$. Imitando il procedimento usato nell’Esempio 4.1 troviamo

$$y_1 - y_0 = f(x_1) - f(x_0) = \Delta f = m\Delta x = m(x_1 - x_0) ,$$

per cui

$$m = \frac{y_1 - y_0}{x_1 - x_0} = \frac{\Delta f}{\Delta x} .$$

Osservazione 4.3 Ovviamente stiamo supponendo che $x_1 \neq x_0$, in quanto altrimenti P_0 e P_1 non potrebbero (perché?) essere due punti del grafico di una sola funzione (a meno che non siano uguali, nel qual caso striglia il tuo assistente e imponigli di misurare due coppie di dati diverse, se vuole sperare di ottenere un qualche risultato).

Una volta trovato m, è facile recuperare anche d: infatti

$$d = f(x_0) - mx_0 = y_0 - mx_0 .$$

Nota che

$$\begin{aligned} f(x_1) - mx_1 &= f(x_0) + \big(f(x_1) - f(x_0)\big) - mx_0 - m(x_1 - x_0) \\ &= f(x_0) - mx_0 + \Delta f - m\Delta x = f(x_0) - mx_0 , \end{aligned}$$

per cui si ottiene lo stesso valore di d sia usando P_0 sia usando P_1.

Osservazione 4.4 Dalla formula $\Delta f = m\Delta x$ possiamo dedurre che

$$f(x) = f\bigl(x_0 + \Delta x\bigr) = f(x_0) + m(x - x_0) = f(x_0) + \frac{\Delta f}{\Delta x} \cdot (x - x_0) \ ,$$

formula che ritroveremo in un contesto diverso nel prossimo capitolo.

Dunque ci basta conoscere due punti del grafico di una funzione lineare per ricavare l'espressione della funzione. Viceversa, data la funzione è facile tracciarne il grafico. Infatti, sappiamo già che il grafico di $f(x) = mx+d$ dev'essere una retta; quindi ci basta trovarne due punti. Per esempio, un punto può essere l'intersezione con l'asse delle ordinate: ponendo $x = 0$ troviamo il punto[2] $(0, d)$. Una volta ottenuto un punto $P_0 = (x_0, y_0)$ del grafico, ogni altro punto $P_1 = (x_1, y_1)$ si ottiene con la formula

$$(x_1, y_1) = P_0 + \bigl(\Delta x, m\Delta x\bigr) \ ,$$

dove $\Delta x = x_1 - x_0$.

Osservazione 4.5 I grafici delle funzioni lineari sono tutte e sole le rette non parallele all'asse delle ordinate. Per avere tutte le rette dobbiamo considerare gli insiemi di equazione $ax + by + c = 0$. Quando $b \neq 0$ ricaviamo $y = -(a/b)x - (c/b)$, cioè il grafico della funzione lineare $f(x) = mx + d$ con $m = -a/b$ e $d = -c/b$. Se invece $b = 0$ (e $a \neq 0$) otteniamo $x = -c/a$, per cui è la retta parallela all'asse delle ordinate passante per il punto $(-c/a, 0)$. Analogamente, se $a = 0$ e $b \neq 0$ otteniamo $y = -c/b$, che è la retta parallela all'asse delle ascisse passante per il punto $(0, -c/b)$ — ovvero il grafico della funzione costante $f(x) = -c/b$.

Nell'Esempio 4.2, oltre a ricavare l'ordinata conoscendo l'ascissa (la percentuale conoscendo la temperatura), abbiamo risolto il problema inverso di trovare l'ascissa conoscendo l'ordinata (la tenperatura conoscendo la percentuale). In altre parole, dato il valore y_0 abbiamo risolto l'equazione lineare $f(x) = y_0$. Siccome $f(x) = mx + d$, vediamo subito che:

- se $m \neq 0$ l'equazione $f(x) = y_0$ ha come unica soluzione $x = (y_0 - d)/m$;
- se $m = 0$ e $d \neq y_0$ l'equazione $f(x) = y_0$ non ha soluzioni;
- se $m = 0$ e $d = y_0$ l'equazione $f(x) = y_0$ ha infinite soluzioni (ogni valore di x va bene).

Osservazione 4.6 Se $m \neq 0$, dire che l'equazione $f(x) = y_0$ ha un'unica soluzione quale che sia $y_0 \in \mathbb{R}$ equivale a dire che la funzione $f\colon \mathbb{R} \to \mathbb{R}$ data da $f(x) = mx+d$ è invertibile. L'inversa è la funzione $f^{-1}\colon \mathbb{R} \to \mathbb{R}$ che fornisce la soluzione dell'equazione: $f^{-1}(y) = (y - d)/m$.

[2] Per questo motivo d è a volte chiamato *intercetta delle ordinate*. Invece, il coefficiente m è spesso chiamato *coefficiente angolare*, per motivi che vedremo nella Sezione 4.10.

La relazione $\Delta f = m\Delta x$ permette di determinare facilmente quando una funzione lineare è crescente o decrescente. In generale, una funzione è *crescente* se aumentando il valore dell'argomento aumenta anche il valore della funzione; ed è *decrescente* se invece aumentando il valore dell'argomento il valore della funzione diminuisce. In altre parole, f è crescente se $x_0 \leq x_1$ implica $f(x_0) \leq f(x_1)$, mentre è decrescente se $x_0 \leq x_1$ implica $f(x_0) \geq f(x_1)$. In altre parole ancora, il grafico di una funzione crescente sale andando verso destra; quello di una funzione decrescente invece scende.

Osservazione 4.7 Una funzione è invece *strettamente crescente* se $x_0 < x_1$ implica $f(x_0) < f(x_1)$, escludendo la possibilità che si abbia $f(x_0) = f(x_1)$; ed è *strettamente decrescente* se $x_0 < x_1$ implica $f(x_0) > f(x_1)$. Infine una funzione crescente o decrescente si dice *monotòna* (e non monòtona, anche se l'idea è la stessa: è una funzione che non cambia mai modo di crescere).

Ora, dire che $x_0 < x_1$ equivale a dire che $\Delta x = x_1 - x_0 > 0$; analogamente, dire che $f(x_0) < f(x_1)$ equivale a dire che $\Delta f = f(x_1) - f(x_0) > 0$ (rispettivamente, $f(x_0) > f(x_1)$ equivale a $\Delta f < 0$). Se f è una funzione lineare, sappiamo che $\Delta f = m\Delta x$; quindi

- se $m > 0$ allora $\Delta x > 0$ implica $\Delta f > 0$, cioè f è strettamente crescente;
- se $m < 0$ allora $\Delta x > 0$ implica $\Delta f < 0$, cioè f è strettamente decrescente;
- se $m = 0$ allora $\Delta f \equiv 0$, cioè f è costante.

Conoscere la crescenza o la decrescenza di una funzione aiuta a trovarne i punti di massimo e di minimo. Diremo che un punto x_0 è un *punto di massimo* (rispettivamente, *punto di minimo*) per una funzione f su un intervallo $[a, b]$ se $f(x_0) \geq f(x)$ per ogni $x \in [a, b]$ (rispettivamente, $f(x_0) \leq f(x)$ per ogni $x \in [a, b]$). In altre parole, x_0 è un punto di massimo (minimo) se $\big(x_0, f(x_0)\big)$ è un punto del grafico di f sopra l'intervallo $[a, b]$ con l'ordinata più alta (bassa). Il valore assunto dalla funzione (l'ordinata del grafico) in un punto di minimo (rispettivamente, di massimo) sull'intervallo $[a, b]$ viene detto *(valore) minimo* (rispettivamente *massimo*) di f sull'intervallo, e viene indicato con $\min f$ (rispettivamente, $\max f$), o con $\min\limits_{x\in[a,b]} f(x)$ (rispettivamente, $\max\limits_{x\in[a,b]} f(x)$) in caso sia importante ricordare l'intervallo che si sta considerando.

Osservazione 4.8 Trovare i punti di massimo o di minimo è fondamentale per le applicazioni della matematica. Infatti, in natura vale spesso un principio del minimo sforzo: la configurazione che si realizza (fra le infinite possibili) è quella che minimizza una qualche quantità. Per esempio, la luce segue il cammino più breve, i semi dei fiori cercano di disporsi in modo da minimizzare lo spreco di spazio, e così via.

Vogliamo trovare minimo e massimo di una funzione f monotona sull'intervallo $[a, b]$. Per definizione di intervallo, abbiamo $a \leq x \leq b$ per ogni $x \in [a, b]$. Se f è crescente, questo implica che $f(a) \leq f(x) \leq f(b)$; quindi *se f è crescente sull'intervallo $[a, b]$ un punto di minimo è a, con valore minimo $f(a)$, e un punto* *di massimo è b, con valore massimo $f(b)$.* Un ragionamento analogo (controlla) ci dice che *se f è decrescente sull'intervallo $[a,b]$ un punto di minimo è b, con valore minimo $f(b)$, e un punto di massimo è a, con valore massimo $f(a)$.*

Osservazione 4.9 Se una funzione è strettamente crescente o decrescente (cioè strettamente monotona) su un intervallo chiuso allora ha un unico punto di minimo e un unico punto di massimo (perché?). Se invece non è strettamente monotona, potrebbe averne anche più di uno; per esempio, se f è costante allora tutti i punti sono contemporaneamente sia di massimo sia di minimo. Invece, il valore minimo e il valore massimo su un dato intervallo sono sempre unici (perché?).

CURIOSITÀ 4.1 Attenzione: dimostrare che un punto di minimo o un punto di massimo esiste può essere a volte anche molto complicato. E certe volte potrebbe anche non esistere. Per esempio, la funzione $f(x) = x$ non ha né minimo né massimo sull'intera retta reale $\mathbb{R}$: per quanto grande o quanto piccolo tu scelga $M \in \mathbb{R}$ esistono sempre x_1, $x_2 \in \mathbb{R}$ tali che $f(x_1) < M < f(x_2)$. Lo stesso problema si verifica su intervalli limitati ma non chiusi: la stessa funzione f non ha né minimo né massimo sull'intervallo aperto $(0,1)$ (perché? Ricordati che 0 e 1 *non* appartengono all'intervallo considerato...). Per fortuna, il *Teorema di Weierstrass* assicura che tutte le funzioni continue (che sono la quasi totalità delle funzioni che considereremo in questo corso, e che definiremo nella Curiosità 4.9) hanno sempre almeno un punto di massimo e almeno un punto di minimo su qualsiasi intervallo *chiuso* della retta reale. Un esempio di funzione *non* continua che non ammette né massimo né minimo su un intervallo chiuso è la funzione $f\colon [-1,1] \to \mathbb{R}$ definita da

$$f(x) = \begin{cases} |x| & \text{se } x \neq -1,\ 0,\ 1, \\ 1/2 & \text{se } x = -1,\ 0,\ 1. \end{cases}$$

Prova a tracciarne il grafico.

Abbiamo osservato che le funzioni lineari con coefficiente angolare non nullo sono sempre strettamente monotone; quindi quanto visto ci permette di trovarne massimi e minimi su intervalli chiusi. Per l'esattezza, se $f(x) = mx + d$ si ha

- se $m > 0$ il punto di minimo di f sull'intervallo $[a,b]$ è a, mentre il punto di massimo è b;
- se $m < 0$ il punto di minimo di f sull'intervallo $[a,b]$ è b, mentre il punto di massimo è a.

Informazioni su crescenza e decrescenza aiutano anche a risolvere le disequazioni. Supponiamo di voler risolvere la disequazione $f(x) \geq y_0$ su un intervallo $[a,b]$ in cui la funzione f sia crescente. Ci sono tre casi possibili:

- se $y_0 \leq \min f = f(a)$, allora $f(x) \geq y_0$ per ogni $x \in [a,b]$, cioè tutti gli $x \in [a,b]$ sono soluzione della disequazione;
- se $y_0 > \max f = f(b)$, allora $f(x) < y_0$ per ogni $x \in [a,b]$, cioè la disequazione non ha soluzione in $[a,b]$;
- se $f(a) = \min f < y_0 \leq \max f = f(b)$ allora $f(x) \geq y_0$ per ogni $x \in [x_0,b]$, dove x_0 è la più piccola soluzione (quando esiste; vedi la Curiosità 4.2) dell'equazione $f(x) = y_0$ in $[a,b]$.

CURIOSITÀ 4.2 Un'altra proprietà non completamente banale delle funzioni continue definite su intervalli chiusi è che per ogni $y_0 \in [\min f, \max f]$ l'equazione $f(x) = y_0$ ammette sempre una soluzione più piccola e una soluzione più grande. Sfortunatamente, le funzioni monotone non sono necessariamente continue; ma c'è un modo per aggirare il problema. Infatti, se f è crescente sull'intervallo $[a,b]$ allora per ogni $y_0 \in [\min f, \max f]$ esiste un più piccolo $x_0 \in [a,b]$ tale che $f(x) \geq y_0$ per ogni $x > x_0$, e $f(x) < y_0$ per ogni $x < x_0$. Se f è continua allora necessariamente $f(x_0) = y_0$, come prima; se f non è continua potrebbe succedere che $f(x_0)$ sia strettamente minore di y_0. Quindi le soluzioni della disequazione $f(x) \geq y_0$ sono gli elementi dell'intervallo chiuso $[x_0, b]$ se $f(x_0) = y_0$, e gli elementi dell'intervallo semiaperto $(x_0, b]$ se $f(x_0) < y_0$. Un esempio di funzione crescente non continua è la $f\colon [-1,1] \to \mathbb{R}$ data da:

$$f(x) = \begin{cases} x & \text{se } -1 \leq x \leq 0, \\ x+1 & \text{se } 0 < x \leq 1. \end{cases}$$

Prova a tracciarne il grafico, e verifica che la disequazione $f(x) \geq 1$ ha come soluzione gli elementi dell'intervallo semiaperto $(0,1]$.

Se vogliamo risolvere la disequazione $f(x) \leq y_0$ su un intervallo $[a,b]$ in cui la funzione f sia crescente, ci sono di nuovo tre casi possibili:

– se $y_0 < \min f = f(a)$, allora $f(x) > y_0$ per ogni $x \in [a,b]$, cioè la disequazione non ha soluzione in $[a,b]$;
– se $y_0 \geq \max f = f(b)$, allora $f(x) \leq y_0$ per ogni $x \in [a,b]$, cioè tutti gli $x \in [a,b]$ sono soluzione della disequazione;
– se $f(a) = \min f \leq y_0 < \max f = f(b)$ allora $f(x) \leq y_0$ per ogni $x \in [a, x_0]$, dove x_0 è la più grande soluzione (quando esiste; vedi la Curiosità 4.?) dell'equazione $f(x) = y_0$ in $[a,b]$.

Osservazione 4.10 Ragionamenti analoghi si applicano al caso di intervalli non chiusi, di intervalli illimitati, alle disequazioni strette (cioè con $>$ o $<$ invece di $\geq$ e $\leq$), e alle funzioni decrescenti; lasciamo il compito di scrivere esplicitamente cosa si ottiene nei vari casi a te e al tuo assistente. *Attenzione:* il tuo obiettivo non dev'essere imparare a memoria tutti i casi possibili, ma capire come si ottengono, in modo da poter ripetere il ragionamento quando ti serve solo nei casi che ti servono (con notevole risparmio di tempo e di memoria).

Vediamo cosa questi ragionamenti ci dicono nel caso delle funzioni lineari. Vogliamo risolvere la disequazione $mx + d \geq y_0$; allora

– Se $m > 0$ (cioè $f(x) = mx + d$ è crescente) allora le soluzioni sono gli elementi della semiretta $[x_0, +\infty)$, dove $x_0 = (y_0 - d)/m = f^{-1}(y_0)$ è l'unica soluzione dell'equazione $mx + d = y_0$.
– Se $m < 0$ (cioè $f(x) = mx+d$ è decrescente) allora le soluzioni sono gli elementi della semiretta $(-\infty, x_0]$, dove $x_0 = (y_0 - d)/m = f^{-1}(y_0)$ è l'unica soluzione dell'equazione $mx + d = y_0$.

In maniera analoga (esercizio per te) si risolve la disequazione $mx + d \leq y_0$.

In particolare, se $m > 0$ la disequazione $mx + d > y_0$ ha soluzione una semiretta della forma $(x_0, +\infty)$ quale che sia $y_0 \in \mathbb{R}$. Questo vuol dire che se $m > 0$ *possiamo rendere $f(x) = mx+d$ arbitrariamente grande a patto di scegliere x sufficientemente grande:* per quanto grande sia $M > 0$ possiamo sempre trovare $x_0 > 0$ (sufficientemente grande) tale che $f(x) > M$ non appena $x > x_0$. In simboli,

$$\forall M > 0 \quad \exists x_0 > 0 : x > x_0 \Longrightarrow f(x) > M \,.$$

Quando questo accade, si dice che $f(x)$ ha *limite* $+\infty$ per x che *tende* a $+\infty$, e si scrive

$$\lim_{x \to +\infty} f(x) = +\infty \; .$$

Sempre supponendo $m > 0$, hai anche visto che la disequazione $mx + d < y_0$ ha soluzione una semiretta della forma $(-\infty, x_0)$ quale che sia $y_0 \in \mathbb{R}$. Questo vuol dire che se $m > 0$ *possiamo rendere* $f(x) = mx + d$ *arbitrariamente negativa a patto di scegliere* x *sufficientemente negativo:* per quanto grande sia $M > 0$ possiamo sempre trovare $x_0 < 0$ (sufficientemente negativo) tale che $f(x) < -M$ non appena $x < x_0$. In simboli,

$$\forall M > 0 \quad \exists x_0 < 0 : \, x < x_0 \Longrightarrow f(x) < -M \; .$$

Stavolta si dice che $f(x)$ ha *limite* $-\infty$ per x che *tende* a $-\infty$, e si scrive

$$\lim_{x \to -\infty} f(x) = -\infty \; .$$

Se $m < 0$ la situazione si inverte. In questo caso la disequazione $mx + d > y_0$ ha soluzione una semiretta della forma $(-\infty, x_0)$ quale che sia $y_0 \in \mathbb{R}$. Questo vuol dire che se $m < 0$ *possiamo rendere* $f(x) = mx + d$ *arbitrariamente grande a patto di scegliere* x *sufficientemente negativo:* per quanto grande sia $M > 0$ possiamo sempre trovare $x_0 < 0$ (sufficientemente negativo) tale che $f(x) > M$ non appena $x < x_0$. In simboli,

$$\forall M > 0 \quad \exists x_0 < 0 : \, x < x_0 \Longrightarrow f(x) > M \; .$$

Si dice che $f(x)$ ha *limite* $+\infty$ per x che *tende* a $-\infty$, e si scrive

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

Infine, se $m > 0$ la disequazione $mx + d < y_0$ ha soluzione una semiretta della forma $(x_0, +\infty)$ quale che sia $y_0 \in \mathbb{R}$. Questo vuol dire che se $m > 0$ *possiamo rendere* $f(x) = mx + d$ *arbitrariamente negativa a patto di scegliere* x *sufficientemente grande:* per quanto grande sia $M > 0$ possiamo sempre trovare $x_0 > 0$ (sufficientemente grande) tale che $f(x) < -M$ non appena $x > x_0$. In simboli,

$$\forall M > 0 \quad \exists x_0 > 0 : \, x > x_0 \Longrightarrow f(x) < -M \; .$$

Stavolta si dice che $f(x)$ ha *limite* $-\infty$ per x che *tende* a $+\infty$, e si scrive

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

Riassumendo, il comportamento di una funzione lineare $f(x) = mx + d$ quando x è sufficientemente grande o sufficientemente negativo è dato da:

– se $m > 0$ allora $\lim\limits_{x \to \pm\infty} = \pm\infty$;

– se $m < 0$ allora $\lim_{x \to \pm\infty} = \mp\infty$.

Con questo abbiamo concluso lo studio delle funzioni lineari. Nelle prossime sezioni cercheremo (per quanto possibile) di studiare in modo analogo funzioni più complesse.

4.2 Funzioni quadratiche

Le funzioni lineari sono tutte monotone: sempre crescenti o decrescenti. Non tutti i fenomeni naturali sono rappresentabili con funzioni monotone; spesso servono funzioni che un po' crescono e un po' decrescono.

Esempio 4.5 Hai provato a far saltare il tuo assistente, sperando che la sua altitudine fosse descritta da una funzione monotona. Invece, sfortunatamente, è salito solo per poco e poi è tornato giù. La sua altitudine è stata inizialmente crescente, ha raggiunto un massimo, e poi è diventata descrescente.

Inoltre, anche le funzioni monotone non è detto che siano lineari, cioè che rappresentino relazioni di proporzionalità.

Esempio 4.6 La superficie esterna di una cellula sferica dipende dal quadrato del raggio della cellula, per cui non aumenta in modo proporzionale al raggio.

Il tipo più semplice di funzioni non monotone (e quindi non lineari) è dato dalle *funzioni quadratiche:* funzioni $f\colon \mathbb{R} \to \mathbb{R}$ della forma

$$f(x) = ax^2 + bx + c\ .$$

Il grafico di una funzione quadratica è una curva chiamata *parabola.*

Curiosità 4.3 Più in generale, una *parabola* nel piano è il luogo dei punti la cui distanza da una retta data (detta *direttrice* della parabola) è uguale alla distanza da un punto dato (detto *fuoco* della parabola). Si può dimostrare che una parabola qualsiasi si ottiene sempre ruotando e traslando il grafico di una funzione quadratica.

Il primo obiettivo di questa sezione è trovare come collegare le proprietà geometriche (l'aspetto) del grafico di una funzione quadratica ai suoi coefficienti. Cominciamo studiando la funzione quadratica più semplice di tutte:

$$f(x) = x^2\ ,$$

il cui grafico è rappresentato nella Figura 4.1.

La prima osservazione evidente è che $f(x) \geq 0$ sempre, e che $f(x) = 0$ se e solo se $x = 0$. In particolare,

(a) $f(x) = x^2$ ha un solo punto di minimo $\overline{x} = 0$, con valore minimo $\overline{y} = f(\overline{x}) = 0$; il punto $(\overline{x}, \overline{y}) = (0, 0)$ è detto *vertice* della parabola grafico di f.

(b) la parabola grafico di $f(x) = x^2$ interseca l'asse delle ordinate nel punto di ordinata $c = f(0) = 0$.

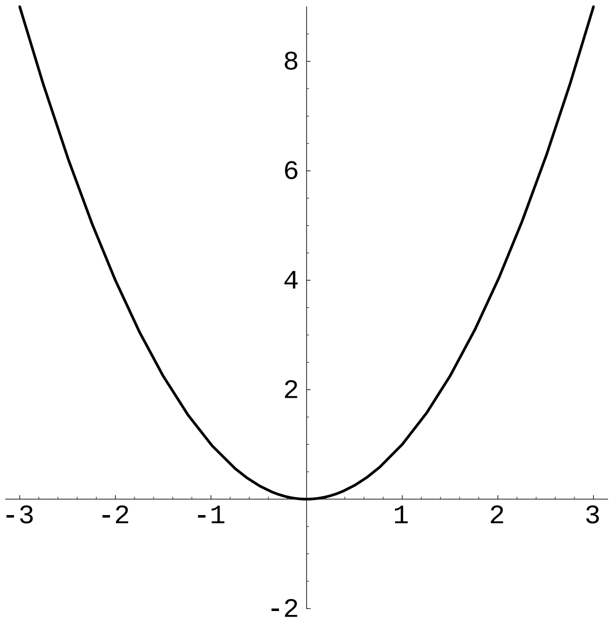

Figura 4.1 $f(x) = x^2$.

Chiaramente, $(-x)^2 = x^2$, cioè $f(-x) = f(x)$ per qualsiasi x. In altri termini,

(c) il grafico di $f(x) = x^2$ è simmetrico rispetto alla retta $x = \overline{x} = 0$ (l'asse delle ordinate), che è detta *asse* della parabola grafico di f.

Osservazione 4.11 Una funzione $f\colon \mathbb{R} \to \mathbb{R}$ tale che $f(-x) = f(x)$ per ogni $x \in \mathbb{R}$ si dice *funzione pari;* se invece $f(-x) = -f(x)$ per ogni $x \in \mathbb{R}$ si dice che f è una *funzione dispari.* Un esempio di funzione dispari è $f(x) = 2x$.

CURIOSITÀ 4.4 Ogni funzione $f\colon \mathbb{R} \to \mathbb{R}$ si può scrivere (in modo unico) come somma di una funzione pari e una funzione dispari. Infatti, ponendo $f_+(x) = \frac{1}{2}\big(f(x) + f(-x)\big)$ e $f_-(x) = \frac{1}{2}\big(f(x) - f(-x)\big)$ si vede subito che f_+ è pari, f_- è dispari, e $f = f_+ + f_-$.

Ora, se $0 \le x_0 < x_1$ abbiamo $f(x_0) = x_0^2 < x_1^2 = f(x_1)$; invece se $x_0 < x_1 \le 0$ abbiamo $0 \le -x_1 < -x_0$ e $f(x_1) = f(-x_1) = (-x_1)^2 < (-x_0)^2 = f(-x_0) = f(x_0)$. Quindi

(d) $f(x) = x^2$ è strettamente decrescente nella semiretta $(-\infty, \overline{x}]$ e strettamente crescente nella semiretta $[\overline{x}, +\infty)$, dove $\overline{x} = 0$. In questo caso, si dice anche che la parabola ha *la concavità rivolta verso l'alto.*

Inoltre, per ogni $y_0 > 0$ la diseguaglianza $f(x) > y_0$ ha come soluzione le semirette $(\sqrt{y_0}, +\infty)$ e $(-\infty, -\sqrt{y_0})$. Quindi possiamo rendere $f(x)$ arbitrariamente grande a patto di prendere x sufficientemente grande o sufficientemente negativo; usando la simbologia dei limiti introdotta nella sezione precedente possiamo dire che

(e) se $f(x) = x^2$ allora $\lim_{x \to \pm\infty} f(x) = +\infty$.

Per concludere la descrizione geometrica della parabola grafico di $f(x) = x^2$ dobbiamo misurarne in qualche modo la larghezza. Un modo per farlo è vedere come cresce l'ordinata allontandosi dal vertice: se cresce molto la parabola è stretta, se cresce poco la parabola è larga. Nel nostro caso si ha

(f) se $f(x) = x^2$ allora $f(x) - f(\overline{x}) = 1 \cdot (x - \overline{x})^2$, dove $\overline{x} = 0$. In particolare, allontanandosi di un'unità dal vertice l'ordinata varia di $a = f(\overline{x}+1) - f(\overline{x}) = 1$.

Possiamo effettuare un'analisi analoga sul grafico della funzione $f(x) = -x^2$. In questo caso si ottiene (vedi la Figura 4.2):

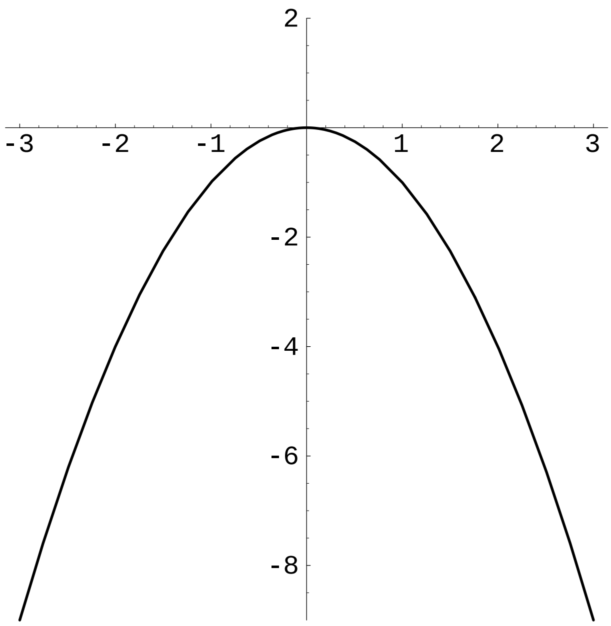

Figura 4.2 $f(x) = -x^2$.

(a) $f(x) = -x^2$ ha un solo punto di massimo $\overline{x} = 0$, e il valore massimo è $\overline{y} = f(\overline{x}) = 0$; il punto $(\overline{x}, \overline{y}) = (0, 0)$ è sempre detto *vertice* della parabola grafico di f.

(b) il grafico di $f(x) = -x^2$ interseca l'asse delle ordinate nel punto di ordinata $c = f(0) = 0$.

(c) il grafico di $f(x) = -x^2$ è simmetrico rispetto alla retta $x = \overline{x} = 0$ (l'asse delle ordinate), che è ancora detta *asse* della parabola grafico di f.

(d) $f(x) = -x^2$ è strettamente crescente nella semiretta $(-\infty, \overline{x}]$ e strettamente decrescente nella semiretta $[\overline{x}, +\infty)$, dove $\overline{x} = 0$. In questo caso, si dice che la parabola ha *la concavità rivolta verso il basso*.

(e) se $f(x) = -x^2$ allora $\lim_{x \to \pm\infty} f(x) = -\infty$.

(f) se $f(x) = -x^2$ allora $f(x) - f(\overline{x}) = -1 \cdot (x - \overline{x})^2$, con $\overline{x} = 0$. In particolare, allontanandosi di un'unità dal vertice l'ordinata varia di $a = f(\overline{x}+1) - f(\overline{x}) = -1$.

Vogliamo far vedere che ogni parabola grafico di funzione quadratica soddisfa opportune variazioni delle proprietà (a)–(f), ed è completamente determinata dalle coordinate $(\overline{x}, \overline{y})$ del vertice e dalla larghezza $a = f(\overline{x}+1) - f(\overline{x})$. Per farlo, vediamo come possiamo spostare il vertice e cambiare la larghezza, e che effetto ha sulla funzione quadratica.

Come primo passo, proviamo a variare la larghezza della parabola. Abbiamo visto che la parabola grafico di x^2 sale di 1 unità se ci spostiamo dal vertice di 1 unità. Se invece salisse di $a > 1$ unità spostandoci dal vertice di 1 unità la parabola sarebbe più stretta (in quanto raggiungiamo l'ordinata 1 prima di $x = 1$); se salisse di $0 < a < 1$ unità spostandoci dal vertice di 1 unità la parabola sarebbe più larga (in quanto raggiungiamo l'ordinata dopo $x = 1$). Per ottenere questo effetto è sufficiente moltiplicare la funzione per a, cioè passare dalla funzione x^2 alla funzione ax^2.

Osservazione 4.12 Questo procedimento si può applicare anche per a negativi. In questo caso $a = -|a| < 0$, per cui moltiplicare per a equivale a moltiplicare prima per $|a| > 0$ (modificando la larghezza della parabola) e poi per -1. Quest'ultima operazione effettua una simmetria rispetto all'asse delle ascisse, ribaltando il grafico; vedi la Figura 4.3.

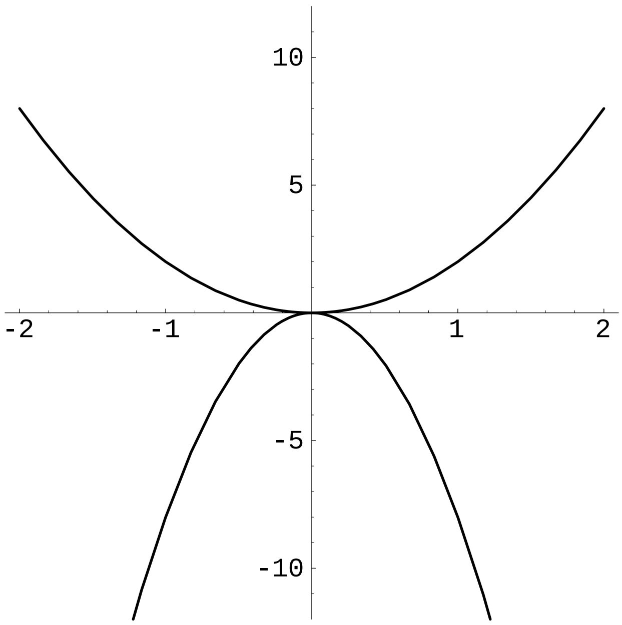

Figura 4.3 $f(x) = 2x^2$ e $f(x) = -8x^2$.

Osservazione 4.13 Moltiplicare per a le ordinate corrisponde a cambiare l'unità di misura (e l'orientazione, se $a < 0$) sull'asse delle ordinate: si ottiene lo stesso effetto dividendo per $|a|$ l'unità di misura (e invertendo l'orientazione se $a < 0$).

Infatti, la vecchia unità di misura, che aveva ordinata 1 nelle vecchie coordinate ora ha ordinata a, per cui la nuova unità di misura (che ha coordinata 1 nelle nuove coordinate) è $1/|a|$ volte la vecchia (con orientazione opposta se $a < 0$).

Le proprietà della funzione $f(x) = ax^2$ si ottengono subito da quelle di x^2 (tenendo presente il segno di a):

(a) $f(x) = ax^2$ ha un solo punto di minimo (se $a > 0$; di massimo se $a < 0$) $\overline{x} = 0$, e il valore minimo (o massimo) è $\overline{y} = f(\overline{x}) = 0$; il vertice della parabola ha ancora coordinate $(\overline{x}, \overline{y}) = (0, 0)$.

(b) Il grafico di $f(x) = ax^2$ interseca l'asse delle ordinate nel punto di ordinata $c = f(0) = 0$.

(c) Il grafico di $f(x) = ax^2$ è simmetrico rispetto all'asse $x = \overline{x} = 0$.

(d) Il grafico di $f(x) = ax^2$ ha la concavità rivolta verso l'alto se $a > 0$, e rivolta verso il basso se $a < 0$.

(e) Se $f(x) = ax^2$ allora $\lim\limits_{x\to\pm\infty} f(x) = +\infty$ se $a > 0$, e $\lim\limits_{x\to\pm\infty} f(x) = -\infty$ se $a < 0$.

(f) Se $f(x) = ax^2$ allora $f(x) - f(\overline{x}) = a \cdot (x - \overline{x})^2$, dove $\overline{x} = 0$. In particolare, allontanandosi di un'unità dal vertice l'ordinata varia di $f(\overline{x} + 1) - f(\overline{x}) = a$.

Proviamo ora a spostare in direzione verticale il vertice della parabola grafico di ax^2. Per portare il vertice nel punto $(0, \gamma)$ è sufficiente traslare l'intero grafico di una distanza pari a γ nella direzione verticale. In altre parole, dobbiamo sommare γ alle ordinate del grafico, cioè passare dalla funzione ax^2 alla funzione $ax^2 + \gamma$; vedi la Figura 4.4.

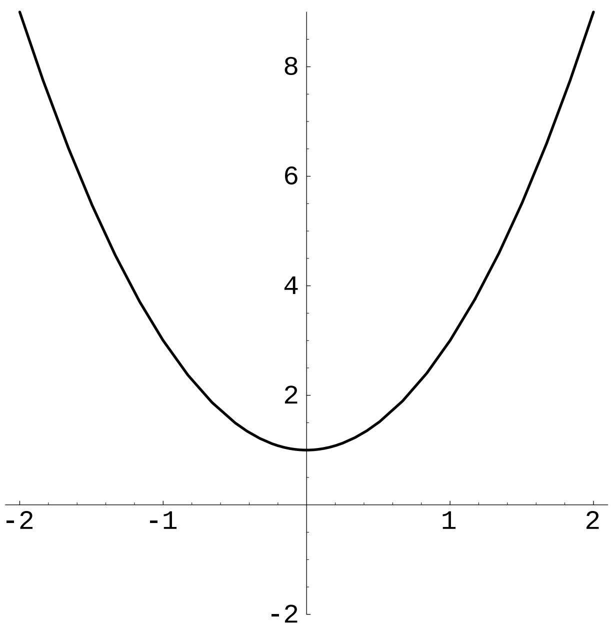

Figura 4.4 $f(x) = ax^2 + \gamma$.

Osservazione 4.14 In generale, traslando verticalmente di una quantità γ il grafico di una funzione f si ottiene il grafico della funzione $f + \gamma$. Inoltre, invece di traslare in direzione verticale il grafico di una quantità γ avremmo potuto traslare in direzione verticale gli assi della quantità $-\gamma$ (cioè sottrarre γ alle ordinate) ottenendo lo stesso risultato. In altre parole, *traslare il piano in direzione verticale di una quantità γ equivale a sottrarre γ alle ordinate.*

Le proprietà della funzione $f(x) = ax^2 + \gamma$ si ottengono subito da quelle di ax^2:

(a) $f(x) = ax^2 + \gamma$ ha un solo punto di minimo (se $a > 0$; di massimo se $a < 0$) $\overline{x} = 0$, e il valore minimo (o massimo) è $\overline{y} = f(\overline{x}) = \gamma$; il vertice della parabola ha ora coordinate $(\overline{x}, \overline{y}) = (0, \gamma)$.
(b) Il grafico di $f(x) = ax^2 + \gamma$ interseca l'asse delle ordinate nel punto di ordinata $c = f(0) = \gamma$.
(c) Il grafico di $f(x) = ax^2 + \gamma$ è simmetrico rispetto all'asse $x = \overline{x} = 0$.
(d) Il grafico di $f(x) = ax^2 + \gamma$ ha la concavità rivolta verso l'alto se $a > 0$, e rivolta verso il basso se $a < 0$.
(e) Se $f(x) = ax^2 + \gamma$ allora $\lim\limits_{x\to\pm\infty} f(x) = +\infty$ se $a > 0$, e $\lim\limits_{x\to\pm\infty} f(x) = -\infty$ se $a < 0$.
(f) Se $f(x) = ax^2 + \gamma$ allora $f(x) - f(\overline{x}) = a \cdot (x - \overline{x})^2$, dove $\overline{x} = 0$. In particolare, allontanandosi di un'unità dal vertice l'ordinata varia di $f(\overline{x} + 1) - f(\overline{x}) = a$.

Il passo successivo consiste nel traslare orizzontalmente il vertice (e quindi l'asse) della parabola. Abbiamo visto che traslare verso l'alto di una quantità γ il grafico è equivalente a sottrarre γ alle ordinate (cioè a spostare gli assi verso il basso di una quantità γ). Per lo stesso motivo, *traslare in direzione orizzontale di una quantità β il grafico equivale a traslare in direzione orizzontale gli assi della quantità $-\beta$, cioè a sottrarre β alle ascisse.* In altre parole, il grafico della funzione $f(x) = a(x-\beta)^2+\gamma$, ottenuta sostituendo $x - \beta$ a x in $ax^2 + \gamma$, è ottenuto traslando orizzontalmente di una quantità β il grafico di $ax^2 + \gamma$. Ne segue che (vedi la Figura 4.5)

(a) $f(x) = a(x - \beta)^2 + \gamma$ ha un solo punto di minimo (se $a > 0$; di massimo se $a < 0$) $\overline{x} = \beta$, e il valore minimo (o massimo) è $\overline{y} = f(\overline{x}) = \gamma$; il vertice della parabola ha quindi coordinate $(\overline{x}, \overline{y}) = (\beta, \gamma)$.
(b) Il grafico di $f(x) = a(x - \beta)^2 + \gamma$ interseca l'asse delle ordinate nel punto di ordinata $c = f(0) = a\beta^2 + \gamma$.
(c) Il grafico di $f(x) = a(x - \beta)^2 + \gamma$ è simmetrico rispetto all'asse $x = \overline{x} = \beta$.
(d) Il grafico di $f(x) = a(x - \beta)^2 + \gamma$ ha la concavità rivolta verso l'alto se $a > 0$, e rivolta verso il basso se $a < 0$.
(e) Se $f(x) = a(x-\beta)^2+\gamma$ allora $\lim\limits_{x\to\pm\infty} f(x) = +\infty$ se $a > 0$, e $\lim\limits_{x\to\pm\infty} f(x) = -\infty$ se $a < 0$.
(f) Se $f(x) = a(x - \beta)^2 + \gamma$ allora $f(x) - f(\overline{x}) = a \cdot (x - \overline{x})^2$, con $\overline{x} = \beta$. Quindi allontanandosi di un'unità dal vertice l'ordinata varia di $f(\overline{x} + 1) - f(\overline{x}) = a$.

Osservazione 4.15 Vale la pena notare che c'è un'altra operazione ancora che potremmo a priori fare: moltiplicare le ascisse per un valore $\alpha \neq 0$, che equivale (come nel caso delle ordinate) a dividere per $|\alpha|$ l'unità di misura sull'asse delle ascisse (e a invertire l'orientazione se $\alpha < 0$). In questo modo arriveremmo a una funzione

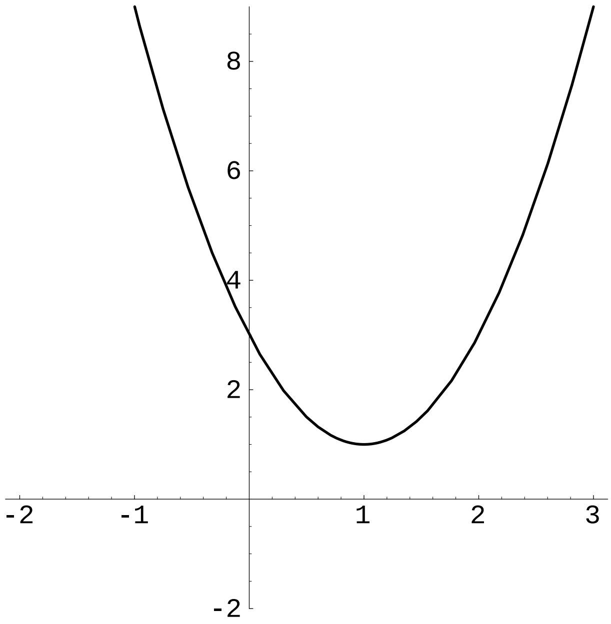

Figura 4.5 $f(x) = a(x-\beta)^2 + \gamma$.

della forma $a(\alpha x - \beta)^2 + \gamma$. Vedremo però fra un attimo che, per le funzioni quadratiche, questa operazione non è necessaria: abbiamo già ottenuto tutte le funzioni quadratiche possibile senza bisogno di ulteriori operazioni. Invece potrebbe essere utile per studiare funzioni più complicate.

Ora, $f(x) = a(x-\beta)^2 + \gamma$ è chiaramente una funzione quadratica: infatti svolgendo il quadrato troviamo

$$a(x-\beta)^2 + \gamma = ax^2 - 2a\beta x + a\beta^2 + \gamma = ax^2 + bx + c \,,$$

dove

$$b = -2a\beta \,, \qquad c = a\beta^2 + \gamma \,. \tag{4.3}$$

La cosa interessante è che vale anche il viceversa: ogni funzione quadratica si può scrivere nella forma $a(x-\beta)^2 + \gamma$. Infatti, se, invertendo le (4.3) (e supponendo ovviamente $a \neq 0$), poniamo

$$\beta = -\frac{b}{2a} \,, \qquad \gamma = c - a\beta^2 = c - \frac{b^2}{4a} = \frac{4ac - b^2}{4a} \,, \tag{4.4}$$

otteniamo

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 + c - \frac{b^2}{4a} = a(x-\beta)^2 + \gamma \,. \tag{4.5}$$

I risultati che abbiamo ottenuto sul grafico delle funzioni della forma $a(x-\beta)^2+\gamma$ si possono quindi tradurre nel caso di funzioni quadratiche qualsiasi:

(a) $f(x) = ax^2 + bx + c$ ha un solo punto di minimo (se $a > 0$; di massimo se $a < 0$) $\overline{x} = -b/2a$, e il valore minimo (o massimo) è $\overline{y} = f(\overline{x}) = c - b^2/4a$; il vertice della parabola ha quindi coordinate $(\overline{x}, \overline{y}) = (-b/2a, c - b^2/4a)$.

(b) Il grafico di $f(x) = ax^2 + bx + c$ interseca l'asse delle ordinate nel punto di ordinata $f(0) = c$.

(c) Il grafico di $f(x) = ax^2 + bx + c$ è simmetrico rispetto all'asse $x = \overline{x} = -b/2a$.

(d) Il grafico di $f(x) = ax^2 + bx + c$ ha la concavità rivolta verso l'alto se $a > 0$, e rivolta verso il basso se $a < 0$. In altre parole, se $a > 0$ la funzione è strettamente decrescente nella semiretta $(-\infty, -b/2a]$ e strettamente crescente nella semiretta $[-b/2a, +\infty)$, mentre se $a < 0$ la funzione è strettamente crescente nella semiretta $(-\infty, -b/2a]$ e strettamente decrescente nella semiretta $[-b/2a, +\infty)$.

(e) Se $f(x) = ax^2 + bx + c$ allora $\lim\limits_{x\to\pm\infty} f(x) = +\infty$ se $a > 0$, e $\lim\limits_{x\to\pm\infty} f(x) = -\infty$ se $a < 0$.

(f) Se $f(x) = ax^2 + bx + c$ allora $f(x) - f(\overline{x}) = a \cdot (x - \overline{x})^2$, dove $\overline{x} = -b/2a$. In particolare, $f(\overline{x} + 1) - f(\overline{x}) = a$.

Dunque data la formula ora siamo in grado di tracciare il grafico. Vediamo ora come risolvere il problema inverso: dato il grafico (o, almeno, alcuni punti del grafico) ricavare la formula.

Un primo caso è quando abbiamo le coordinate $(\overline{x}, \overline{y})$ del vertice e (supponendo che il vertice non sia sull'asse delle ordinate, cioè che $\overline{x} \neq 0$) il punto $(0, c)$ di intersezione del grafico con l'asse delle ordinate. Allora i conti precedenti, e in particolare le (4.3), ci dicono che la funzione dev'essere $f(x) = ax^2 + bx + c$ con a e b dati da

$$b = -2a\overline{x}\,, \qquad a = \frac{c - \overline{y}}{\overline{x}^2}\,.$$

Se invece il vertice è sull'asse delle ordinate, cioè $\overline{x} = 0$, allora sappiamo soltanto che $f(x) = ax^2 + c$ con $c = \overline{y}$; per trovare a servono altre informazioni (quali, per esempio, la larghezza della parabola).

Spesso, invece, conosciamo alcuni punti del grafico, senza però sapere quale sia il vertice. Per determinare la funzione, servono tre punti; vediamo come in un esempio.

Esempio 4.7 *Torniamo a studiare i semi di pomodoro dell'Esempio 4.1. Sai già che alla temperatura di* 12 °C *germoglia il* 40% *dei semi, mentre alla temperatura di* 15 °C *germoglia il* 70% *dei semi. Un'ulteriore misurazione ha rivelato che alla temperatura di* 9 °C *germoglia il* 20% *dei semi. Dimostra che allora la relazione fra la temperatura e la percentuale di semi che gemogliano non può essere lineare. Supponendo che sia quadratica, determinala.* Indichiamo nuovamente con $P(T)$ la percentuale di semi che germogliano alla temperatura T. Noi sappiamo che $P(9) = 20$, $P(12) = 40$ e $P(15) = 70$. Se P fosse una funzione lineare, $\Delta P/\Delta T$

dovrebbe essere costante; invece

$$\frac{P(15)-P(12)}{15-12}=\frac{70-40}{3}=10\neq\frac{20}{3}=\frac{40-20}{3}=\frac{P(12)-P(9)}{12-9}\;.$$

Supponiamo allora che $P\colon\mathbb{R}\to\mathbb{R}$ sia una funzione quadratica, cioè che si abbia $P(T)=aT^2+bT+c$; dobbiamo trovare a, b, $c\in\mathbb{R}$ in modo da avere $P(9)=20$, $P(12)=40$ e $P(15)=70$. In altre parole, a, b e c devono soddisfare il seguente sistema di equazioni lineari:

$$\begin{cases} 81a+9b+c=P(9)=20\;,\\ 144a+12b+c=P(12)=40\;,\\ 225a+15b+c=P(15)=70\;. \end{cases} \tag{4.6}$$

Per risolvere questo sistema, sottraiamo la prima equazione dalla seconda, e la seconda dalla terza; otteniamo

$$\begin{cases} 63a+3b=20\;,\\ 81a+3b=30\;. \end{cases} \tag{4.7}$$

Sottraendo di nuovo la prima equazione dalla seconda otteniamo

$$18a=10\;,\quad\text{cioè}\quad a=\frac{10}{18}=\frac{5}{9}\;.$$

Sostituendo questo valore nella seconda equazione in (4.7) troviamo

$$81\cdot\frac{5}{9}+3b=20\;,\quad\text{cioè}\quad b=-5\;;$$

e sostituendo i valori di a e b trovati nella prima equazione in (4.6) recuperiamo infine

$$81\cdot\frac{5}{9}-9\cdot5+c=20\;,\quad\text{cioè}\quad c=20\;.$$

Quindi la formula cercata è

$$P(T)=\frac{5}{9}T^2-5T+20\;.$$

Questa formula ha qualche vantaggio su quella lineare. Per esempio, non è mai negativa; infatti, ha minimo per $T=-(-5)/2(5/9)=9/2$, con valore minimo $P(9/2)=20-(-5)^2/4(5/9)=35/4>0$. Ma anche lei può essere valida solo in un determinato intervallo di temperature. Infatti, $P(T)$ ricomincia ad aumentare quando la temperatura scende sotto $9/2$ °C, comportamento biologicamente alquanto improbabile; e $P(T)>100$ se T è troppo grande (o sufficientemente negativo). Per esempio, $P(18)=110$.

Il procedimento usato nel precedente esempio può essere applicato a qualsiasi funzione quadratica. Supponiamo di voler trovare la funzione quadratica $f(x) = ax^2 + bx + c$ il cui grafico passi per i punti (x_1, y_1), (x_2, y_2) e (x_3, y_3), con x_1, x_2 ed x_3 tutti distinti. Vogliamo quindi trovare a, b e c in modo che $f(x_j) = y_j$ per $j = 1, 2, 3$. In altre parole, a, b e c devono risolvere il sistema

$$\begin{cases} x_1^2 a + x_1 b + c = y_1 \ , \\ x_2^2 a + x_2 b + c = y_2 \ , \\ x_3^2 a + x_3 b + c = y_3 \ . \end{cases}$$

Sottraiamo la prima equazione dalla seconda, e la seconda dalla terza; otteniamo

$$\begin{cases} (x_2^2 - x_1^2)a + (x_2 - x_1)b = y_2 - y_1 \ , \\ (x_3^2 - x_2^2)a + (x_3 - x_2)b = y_3 - y_2 \ . \end{cases} \tag{4.8}$$

Siccome $x_2^2 - x_1^2 = (x_2 - x_1)(x_2 + x_1)$ e $x_2 - x_1 \neq 0$, possiamo dividere la prima equazione per $x_2 - x_1$. Analogamente possiamo dividere la seconda equazione per $x_3 - x_2$, e otteniamo

$$\begin{cases} (x_2 + x_1)a + b = \frac{y_2 - y_1}{x_2 - x_1} \ , \\ (x_3 + x_2)a + b = \frac{y_3 - y_2}{x_3 - x_2} \ . \end{cases}$$

Sottraendo di nuovo la prima equazione dalla seconda otteniamo

$$(x_3 - x_1)a = \frac{y_3 - y_2}{x_3 - x_2} - \frac{y_2 - y_1}{x_2 - x_1} \ .$$

Siccome $x_3 - x_1 \neq 0$, da questa equazione possiamo ricavare a; sostituendo il valore trovato nel sistema precedente otteniamo b, e sostituendo nel sistema iniziale troviamo anche c.

Osservazione 4.16 La generica funzione lineare dipendeva da due parametri (m e d); per determinarla avevamo bisogno di conoscere due punti del grafico. La generica funzione quadratica dipende da tre parametri (a, b e c); per determinarla abbiamo bisogno di conoscere tre punti del grafico. Tutto ciò non è un caso: si può dimostrare che se abbiamo una famiglia di funzioni dipendenti da k parametri, per determinare univocamente una funzione della famiglia servono k condizioni (indipendenti in un senso opportuno), quali il conoscere k punti del grafico. Vedremo un altro esempio di questo fenomeno nella Sezione 4.4.

La (4.5) è molto utile anche per risolvere le equazioni di secondo grado. L'idea è che l'equazione $ax^2 + bx + c = y_0$ ha soluzione se e solo se la retta $y = y_0$ interseca il grafico di $f(x) = ax^2 + bx + c$. Questo accade solo se o $a > 0$ e y_0 è maggiore del minimo di f, oppure $a < 0$ e y_0 è minore del massimo di f. Quindi bisogna confrontare y_0 con l'ordinata $\overline{y} = \gamma$ del vertice, che si legge facilmente da (4.5). Per la precisione, l'equazione $ax^2 + bx + c = y_0$ diventa

$$a(x - \beta)^2 + \gamma = y_0 \ , \quad \text{cioè} \quad (x - \beta)^2 = \frac{y_0 - \gamma}{a} \ .$$

Quindi ha soluzione reale se e solo se $(y_0 - \gamma)/a \geq 0$, e in tal caso le soluzioni sono

$$x_\pm = \beta \pm \sqrt{\frac{y_0 - \gamma}{a}} = \frac{-b \pm \sqrt{b^2 - 4a(c - y_0)}}{2a} \,. \tag{4.9}$$

Il caso $y_0 = 0$ è particolarmente interessante (e tutti gli altri possono esservi ricondotti sostituendo $c - y_0$ al posto di c). La quantità

$$D = b^2 - 4ac = -4a\overline{y}$$

è detta *discriminante* della funzione $f(x) = ax^2 + bx + c$. Il discriminante è positivo se e solo se a e l'ordinata $\overline{y}$ del vertice hanno segno opposto, e si annulla se e solo se l'ordinata del vertice si annulla, cioè se e solo se il vertice è sull'asse delle ascisse. Insomma, o usando la posizione del vertice e la concavità del grafico, oppure usando il segno del discriminante in (4.9), giungiamo alla conclusione che l'equazione $ax^2 + bx + c = 0$ ha

- due soluzioni reali distinte se $D > 0$, cioè se a e $\overline{y}$ hanno segno opposto;
- un'unica soluzione reale se $D = 0$, cioè se $\overline{y} = 0$, e in tal caso la soluzione è $\overline{x}$;
- nessuna soluzione reale se $D < 0$, cioè se a e $\overline{y}$ hanno lo stesso segno.

Come esercizio, riotteni questo risultato tracciando il grafico di $ax^2 + bx + c$ nei vari casi.

Usando le informazioni che abbiamo sulla crescenza e decrescenza delle funzioni quadratiche, possiamo facilmente trovare i massimi e i minimi in intervalli chiusi. Ci sono due casi da considerare:

- se l'intervallo $[a_0, a_1]$ non contiene l'ascissa del vertice di $f(x) = ax^2 + bx + c$, allora f è monotona in quell'intervallo, per cui (come abbiamo visto nella precedente sezione) i punti di minimo e massimo di f in $[a_0, a_1]$ sono gli estremi a_0 e a_1.
- se l'intervallo $[a_0, a_1]$ contine l'ascissa $\overline{x}$ del vertice, sappiamo già che il punto di minimo (se $a > 0$) o di massimo (se $a < 0$) è $\overline{x}$. Siccome f è monotona nei due intervalli $[a_0, \overline{x}]$ e $[\overline{x}, a_1]$, si vede subito (perché?) che il punto di massimo (se $a > 0$) o di minimo (se $a < 0$) è quello[3] fra i due estremi a_0 e a_1 su cui f assume il valore più grande (se $a > 0$) o più piccolo (se $a < 0$).

Queste tecniche ci permettono anche di risolvere facilmente le disequazioni di secondo grado. Perché la disequazione $ax^2 + bx + c \geq y_0$ possa avere soluzione occorre che il grafico di $f(x) = ax^2 + bx + c$ sia in qualche punto al di sopra della retta $y = y_0$. Mettendo insieme (4.9) con ciò che sappiamo sulla crescenza e decrescenza di f troviamo

- se $a > 0$ e
  - $y_0 \leq \overline{y}$, la disequazione $ax^2 + bx + c \geq y_0$ è soddisfatta per tutti i valori di $x \in \mathbb{R}$;

[3] O entrambi se $f(a_0) = f(a_1)$.

– $y_0 > \overline{y}$, la disequazione $ax^2 + bx + c \geq y_0$ è soddisfatta per $x \in (-\infty, x_-]$ e per $x \in [x_+, +\infty)$, dove $x_\pm$ sono dati da (4.9);

- se $a < 0$ e
  – $y_0 > \overline{y}$, la disequazione $ax^2 + bx + c \geq y_0$ non è mai soddisfatta;
  – $y_0 \leq \overline{y}$, la disequazione $ax^2 + bx + c \geq y_0$ è soddisfatta per $x \in [x_+, x_-]$, dove $x_\pm$ sono dati da (4.9).

Analoghi risultati (esercizio: se sei confuso, aiutati tracciando il grafico nei vari casi) si trovano per la disequazione $ax^2 + bx + c \leq y_0$.

Osservazione 4.17 Quando $y_0 = 0$, il segno di $\overline{y}$ è legato al segno di a tramite il segno del discriminante. In particolare, se $a > 0$ abbiamo $\overline{y} \geq 0$ se e solo se $D \leq 0$, mentre se $a < 0$ abbiamo $0 > \overline{y}$ se e solo se $D < 0$. Quindi quando $y = 0$ possiamo riformulare (esercizio per te) i risultati precedenti usando a e D invece di a e $\overline{y}$.

Osservazione 4.18 Come fatto nella scorsa sezione, i risultati sulle disequazioni ci permettono anche di studiare l'andamento all'infinito delle funzioni quadratiche. Per esempio, abbiamo appena visto che se $a > 0$ allora per ogni $y_0 \in \mathbb{R}$ possiamo trovare $x_\pm \in \mathbb{R}$ tali che $f(x) \geq y_0$ non appena $x \geq x_+$ oppure $x \leq x_-$. In altre parole, possiamo rendere $f(x)$ arbitrariamente grande a patto di scegliere x sufficientemente grande o sufficientemente negativo. Usando la terminologia già introdotta, abbiamo quindi dimostrato che

$$a > 0 \qquad \Longrightarrow \qquad \lim_{x \to \pm\infty} ax^2 + bx + c = +\infty \; ,$$

in accordo con quanto avevamo già visto. In maniera analoga si dimostra che

$$a < 0 \qquad \Longrightarrow \qquad \lim_{x \to \pm\infty} ax^2 + bx + c = -\infty \; .$$

4.3 Il metodo dei minimi quadrati

In questa sezione presenteremo due applicazioni dello studio delle funzioni quadratiche che abbiamo appena completato.

La prima applicazione consiste nel mantenere una promessa fatta nell'Osservazione 3.27. Siano $x_1, \ldots, x_n \in \mathbb{R}$ dei numeri reali (dei dati); vogliamo trovare il punto di minimo della funzione

$$f(x) = \sum_{i=1}^{n} (x - x_i)^2 \; .$$

Se sviluppiamo i quadrati, vediamo subito che f è una funzione quadratica:

$$f(x) = \sum_{i=1}^{n} (x^2 - 2x_i x + x_i^2) = nx^2 - 2\left(\sum_{i=1}^{n} x_i\right) x + \sum_{i=1}^{n} x_i^2 \; ,$$

per cui $f(x) = ax^2 + bx + c$ con

$$a = n \;, \qquad b = -2\sum_{i=1}^{n} x_i \;, \qquad c = \sum_{i=1}^{n} x_i^2 \;.$$

Quindi il punto di minimo di f è

$$\overline{x} = -\frac{b}{2a} = \frac{1}{n}\sum_{i=1}^{n} x_i \;,$$

cioè la media aritmetica dei dati, come promesso.

La seconda applicazione consiste in un metodo (detto *metodo dei minimi quadrati*) per trovare la retta che meglio approssima un dato insieme di dati.

Supponiamo di avere n coppie di dati $(x_1, y_1), \ldots, (x_n, y_n)$, e di sospettare che le ordinate dipendano in modo lineare dalle ascisse. Anche se la nostra ipotesi è corretta, è molto improbabile che le n coppie di dati giacciano esattamente su una retta, in quanto non possiamo evitare gli errori sperimentali; abbiamo quindi bisogno di una tecnica che ci fornisca la “migliore” (in un senso da specificare) approssimazione lineare di questi dati, e al contempo una misura della bontà di questa approssimazione — in quanto, se la “migliore” approssimazione fosse cattiva vorrebbe dire che la nostra ipotesi di dipendenza lineare non è compatibile con i dati, e quindi dev’essere scartata.

Cominciamo con definire quanto il grafico di una funzione lineare $f(x) = mx + d$ approssima l’insieme $\{(x_1, y_1), \ldots, (x_n, y_n)\}$. La coppia (x_i, y_i) appartiene al grafico di f se e solo se $y_i = mx_i + d$; quindi l’errore $\delta_i = mx_i + d - y_i$ misura la distanza che c’è fra il dato sperimentale (x_i, y_i) e il dato teorico $\bigl(x_i, f(x_i)\bigr)$ che si avrebbe se la funzione f rappresentasse esattamente il fenomeno che stiamo studiando. Abbiamo quindi n errori, $\delta_1, \ldots, \delta_n$; tenendo presente che a noi non importa il segno dell’errore ma solo la sua grandezza, e ricordando quanto fatto studiando la varianza, una misura di quanto la funzione $f(x) = mx + d$ approssima i dati è data dalla media dei quadrati degli errori:

$$S(m, d) = \frac{1}{n}\sum_{i=1}^{n} \delta_i^2 = \frac{1}{n}\sum_{i=1}^{n} (mx_i + d - y_i)^2 \;.$$

Nota che la funzione S dipende dai due parametri m e d che determinano la funzione lineare f; quindi S è una funzione di due variabili reali.

Il nostro obiettivo è trovare il punto di minimo di S, cioè i valori $\overline{m}$ e $\overline{d}$ di m e d che rendono $S(m, d)$ più piccola possibile[4]. La corrispondente funzione lineare $f(x) = \overline{m}x + \overline{d}$ sarà allora quella che meglio approssima i dati da cui siamo partiti; e ci rimarrà solo da trovare una misura di quanto buona sia questa approssimazione.

Per trovare questo punto di minimo procederemo in questo modo. Prima di tutto faremo vedere che, per ogni $m \in \mathbb{R}$ fissato, la funzione $d \mapsto S(m, d)$ è una

[4] Ed è questo il motivo per cui questo metodo si chiama dei *minimi* quadrati.

funzione quadratica di d con coefficiente del termine quadrato positivo; quindi ammette un unico punto di minimo, che indicheremo con $d_0(m)$; vedremo anche come dipende da m. Poi dimostreremo che anche la funzione $S\big(m, d_0(m)\big)$ è una funzione quadratica (di m, stavolta) con coefficiente del termine quadrato positivo; quindi anche lei ha un unico punto di minimo $\overline{m}$, a cui corrisponde il valore $\overline{d} = d_0(\overline{m})$. Allora $(\overline{m}, \overline{d})$ è il punto di minimo cercato. Infatti, per ogni $(m, d) \in \mathbb{R}^2$ abbiamo

$$S(m,d) \geq S\big(m, d_0(m)\big) \geq S\big(\overline{m}, d_0(\overline{m})\big) = S(\overline{m}, \overline{d})$$

come voluto (studia bene la precedente catena di disuguaglianze fin quando non sei certo d'aver capito perché è vera e perché è proprio quello che ci serve).

Ok, cominciamo. Sviluppando i quadrati nella definizione di $S(m, d)$ otteniamo

$$\begin{aligned} S(m,d) &= \frac{1}{n}\sum_{i=1}^{n}\big(x_i^2 m^2 + d^2 + y_i^2 + 2x_i m d - 2x_i y_i m - 2y_i d\big) \\ &= d^2 + 2(m\overline{x} - \overline{y})d + m^2\,\frac{1}{n}\sum_{i=1}^{n} x_i^2 - 2m\,\frac{1}{n}\sum_{i=1}^{n} x_i y_i + \frac{1}{n}\sum_{i=1}^{n} y_i^2\ , \end{aligned}$$

dove $\overline{x}$ è la media aritmetica di $x_1, \ldots, x_n$ e $\overline{y}$ è la media aritmetica di $y_1, \ldots, y_n$. Quindi per ogni m fissato $d \mapsto S(m, D)$ è effettivamente una funzione quadratica di d, in quanto possiamo scrivere $S(m, d) = ad^2 + bd + c$ con

$$a = 1\ , \quad b = 2(m\overline{x} - \overline{y})\ , \quad c = m^2\,\frac{1}{n}\sum_{i=1}^{n} x_i^2 - 2m\,\frac{1}{n}\sum_{i=1}^{n} x_i y_i + \frac{1}{n}\sum_{i=1}^{n} y_i^2\ .$$

In particolare, $a = 1 > 0$ e il punto di minimo $d_0(m)$ di questa funzione è

$$d_0(m) = -\frac{b}{2a} = \overline{y} - m\overline{x}\ .$$

Per calcolare $S\big(m, d_0(m)\big)$ ricordiamoci che il valore minimo di una funzione quadratica (con $a > 0$) è $c - b^2/4a$, per cui

$$\begin{aligned} S\big(m, d_0(m)\big) &= m^2\,\frac{1}{n}\sum_{i=1}^{n} x_i^2 - 2m\,\frac{1}{n}\sum_{i=1}^{n} x_i y_i + \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \frac{1}{4}4(m\overline{x} - \overline{y})^2 \\ &= \left[\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2\right] m^2 - 2\left[\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x}\cdot\overline{y}\right] m + \frac{1}{n}\sum_{i=1}^{m} y_i^2 - \overline{y}^2 \\ &= \tilde{a}m^2 + \tilde{b}m + \tilde{c}\ . \end{aligned}$$

Dunque abbiamo ottenuto, come promesso, una funzione quadratica di m, con

$$\tilde{a} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2\ , \quad \tilde{b} = -2\left[\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x}\cdot\overline{y}\right]\ , \quad \tilde{c} = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \overline{y}^2\ .$$

Inoltre, il coefficiente $\tilde{a}$ del termine quadrato è positivo; infatti, l'Osservazione 3.32 ci dice che

$$\tilde{a} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2 = \mathrm{Media}(x_i^2) - \mathrm{Media}(x_i)^2 = \mathrm{Var}(x_i) > 0 \ ,$$

come voluto[5].

Quindi la funzione $S\big(m, d_0(m)\big)$ ammette un unico punto di minimo

$$\overline{m} = -\frac{\tilde{b}}{2\tilde{a}} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x}\cdot\overline{y}}{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2} \ , \tag{4.10}$$

e la funzione $S(m,d)$ ammette un unico punto di minimo $(\overline{m}, \overline{d})$ con

$$\overline{d} = \overline{y} - \overline{m}\,\overline{x} \ . \tag{4.11}$$

La retta grafico della funzione $f(x) = \overline{m}x + \overline{d}$ che meglio approssima i dati si chiama *retta di regressione* (*lineare*).

Osservazione 4.19 Mentre il modo migliore di calcolare $\overline{d}$ è usare la formula (4.11), ci sono altre formule per il calcolo di $\overline{m}$ oltre a (4.10). Prima di tutto notiamo che possiamo scrivere (4.10) così:

$$\overline{m} = \frac{\mathrm{Media}(x_i y_i) - \mathrm{Media}(x_i)\mathrm{Media}(y_i)}{\mathrm{Media}(x_i^2) - \mathrm{Media}(x_i)^2} \ , \tag{4.12}$$

o anche come

$$\overline{m} = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{\overline{x^2} - \overline{x}^2} \ , \tag{4.13}$$

dove $\overline{xy}$ è la media aritmetica dei prodotti $x_1y_1, \ldots, x_ny_n$, e $\overline{x^2}$ è la media aritmetica dei quadrati $x_1^2, \ldots, x_n^2$, per cui $\overline{m}$ *è la differenza fra la media dei prodotti e il prodotto delle medie divisa per la differenza fra la media dei quadrati e il quadrato della media.*

Un'altra formula si ottiene notando che, come abbiamo già visto, il denominatore di (4.10) è uguale a $\mathrm{Var}(x_i) = \frac{1}{n}\sum_{i=1}^{n}(\overline{x} - x_i)^2$. Una formula analoga vale anche per il numeratore: infatti

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x}\cdot\overline{y} &= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x}\,\frac{1}{n}\sum_{i=1}^{n} y_i - \overline{y}\,\frac{1}{n}\sum_{i=1}^{n} x_i + \overline{x}\cdot\overline{y} \\ &= \frac{1}{n}\sum_{i=1}^{n}(\overline{x} - x_i)(\overline{y} - y_i) \ ,\end{aligned}$$

[5] La varianza non può essere nulla, a meno che tutti gli x_i siano uguali; ma in tal caso sapremmo già che i dati giacciono su una retta (verticale), e quindi non avremmo neppure cominciato questi conti.

per cui possiamo scrivere

$$\overline{m} = \frac{\sum_{i=1}^{n}(\overline{x} - x_i)(\overline{y} - y_i)}{\sum_{i=1}^{n}(\overline{x} - x_i)^2} \ . \tag{4.14}$$

Quale formula usare dipende dalle situazioni, da quali altri calcoli hai già effettuato o devi effettuare, da quale ti ricordi meglio e anche dai tuoi gusti personali.

Rimane da stabilire quanto bene la retta di regressione approssima i dati. Una prima informazione ci è data dal valore minimo che abbiamo trovato,

$$\begin{aligned} S(\overline{m}, \overline{d}) &= \tilde{c} - \frac{\tilde{b}^2}{4\tilde{a}} = \tilde{c} - \tilde{a}\,\overline{m}^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \overline{y}^2 - \frac{\left[\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \overline{x}\cdot\overline{y}\right]^2}{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2} \\ &= \frac{(\overline{y^2} - \overline{y}^2)(\overline{x^2} - \overline{x}^2) - (\overline{xy} - \overline{x}\cdot\overline{y})^2}{\overline{x^2} - \overline{x}^2} \ , \end{aligned}$$

dove $\overline{y^2}$ è la media aritmetica di $y_1^2, \dots, y_n^2$, come al solito. Ora, le coppie di dati stanno tutte sulla retta di regressione se e solo se $S(\overline{m}, \overline{d}) = 0$. Però il valore di $S(\overline{m}, \overline{d})$ da solo non è una buona misura della qualità della retta di regressione, in quanto ha il solito problema degli errori assoluti: se i dati sono grandi allora l'errore è grande in valore assoluto, anche quando è piccolo rispetto ai valori assoluti dei dati. Ci serve invece un errore relativo. Siccome $S(\overline{m}, \overline{d})$ misura la media degli errori quadratici nelle ordinate, la quantità giusta a cui confrontarla è lo scarto quadratico medio delle ordinate, cioè la varianza[6] $\mathrm{Var}(y_i) = \overline{y^2} - \overline{y}^2$. Quindi siamo condotti a considerare la quantità

$$\frac{S(\overline{m}, \overline{d})}{\mathrm{Var}(y_i)} = 1 - \frac{(\overline{xy} - \overline{x}\cdot\overline{y})^2}{(\overline{x^2} - \overline{x}^2)(\overline{y^2} - \overline{y}^2)} \ ,$$

che è sempre maggiore o uguale di zero, e si annulla se e solo se tutti i dati sono sulla retta di regressione. Inoltre, è anche sempre minore o uguale di 1 (perché?); quindi anche se i dati sono molto grandi rimane con valore assoluto limitato.

Abbiamo quasi finito: rimane da fare un passaggio analogo a quello che porta dalla varianza alla deviazione standard. La quantità $S(\overline{m}, \overline{d})\big/\mathrm{Var}(y_i)$ si annulla se e solo se $(\overline{xy} - \overline{x}\cdot\overline{y})^2\big/(\overline{x^2} - \overline{x}^2)(\overline{y^2} - \overline{y}^2) = 1$, cioè (estraendo la radice quadrata) se e solo se $(\overline{xy} - \overline{x}\cdot\overline{y})\big/\sqrt{(\overline{x^2} - \overline{x}^2)(\overline{y^2} - \overline{y}^2)} = \pm 1$. Introduciamo allora il *coefficiente di correlazione di Pearson*

$$\mathrm{CP} = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{\sqrt{(\overline{x^2} - \overline{x}^2)(\overline{y^2} - \overline{y}^2)}} \in [-1, 1] \ .$$

[6] Che è nulla se e solo se tutte le y_i sono uguali; ma in tal caso i dati sono chiaramente su una retta (orizzontale), e di nuovo non ci saremmo imbarcati in questi calcoli.

Per quanto abbiamo detto, il coefficiente di correlazione di Pearson misura la bontà dell'approssimazione fornita dalla retta di regressione: se è sufficientemente vicino a 1 o a -1, allora l'approssimazione è buona; se invece è vicino a 0, vuol dire che i dati non seguono affatto un andamento lineare.

Osservazione 4.20 "Sufficientemente vicino a ± 1" di solito vuol dire almeno 0.9 in valore assoluto; almeno 0.95 è anche meglio.

Osservazione 4.21 Il segno del coefficiente di Pearson è lo stesso di $\overline{m}$.

Osservazione 4.22 I conti fatti nell'Osservazione 4.19 ci forniscono un'altra formula per il coefficiente di Pearson:

$$\mathrm{CP} = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\mathrm{DS}(x_i)\mathrm{DS}(y_i)} = \frac{\sum_{i=1}^{n}(\overline{x} - x_i)(\overline{y} - y_i)}{\sqrt{\sum_{i=1}^{n}(\overline{x} - x_i)^2 \sum_{j=1}^{n}(\overline{y} - y_j)^2}} .$$

Concludiamo questa sezione con un esempio di calcolo della retta di regressione e del coefficiente di Pearson; altri esempi li vedremo nella Sezione 4.9.

ESEMPIO 4.8 Riprendiamo il nostro gruppo di 15 cavie; vogliamo vedere se c'è una relazione lineare fra il loro peso (in decigrammi) e la loro età (in giorni; sono cavie molto giovani). Per procedere prepariamo una tabella (Tabella 4.1) con cinque colonne: l'età (la nostra x), il peso (la nostra y), i prodotti xy, i quadrati x^2, e i quadrati y^2. Poi calcoliamo la media aritmetica dei dati di ciascuna colonna; con questi dati possiamo trovare i coefficienti della retta di regressione e il coefficiente di Pearson.

Cavia	*Età* (x)	*Peso* (y)	xy	x^2	y^2
1	61	28	1708	3721	784
2	76	32	2432	5776	1024
3	80	37	2960	6400	1369
4	66	29	1914	4356	841
5	71	31	2201	5041	961
6	68	30	2040	4624	900
7	78	32	2496	6084	1024
8	55	26	1430	3025	676
9	74	32	2368	5476	1024
10	60	27	1620	3600	729
11	65	29	1885	4225	841
12	70	30	2100	4900	900
13	64	28	1792	4096	784
14	73	31	2263	5329	961
15	68	31	2108	4624	961
Media	68.6	30.2	2087.8	4751.8	918.6

TABELLA 4.1

Usando per esempio la formula (4.13) otteniamo

$$\overline{m} = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\overline{x^2} - \overline{x}^2} = \frac{2087.8 - 68.6 \cdot 30.2}{4751.8 - 68.6^2} \simeq 0.351 ,$$
$$\overline{d} = \overline{y} - \overline{m} \cdot \overline{x} \simeq 30.2 - 0.351 \cdot 67 \simeq 6.136 ,$$

per cui la retta di regressione è

$$f(x) = 0.351x + 6.136 .$$

Il coefficiente di Pearson è

$$\mathrm{CP} = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sqrt{(\overline{x^2} - \overline{x}^2)(\overline{y^2} - \overline{y}^2)}} = \frac{2087.8 - 68.6 \cdot 30.2}{\sqrt{(4751.8 - 68.6^2)(918.6 - 30.2^2)}} \simeq 0.927 ,$$

per cui la retta di regressione approssima piuttosto bene i dati, come si può vedere dalla Figura 4.6, che contiene sia i dati sia la retta di regressione.

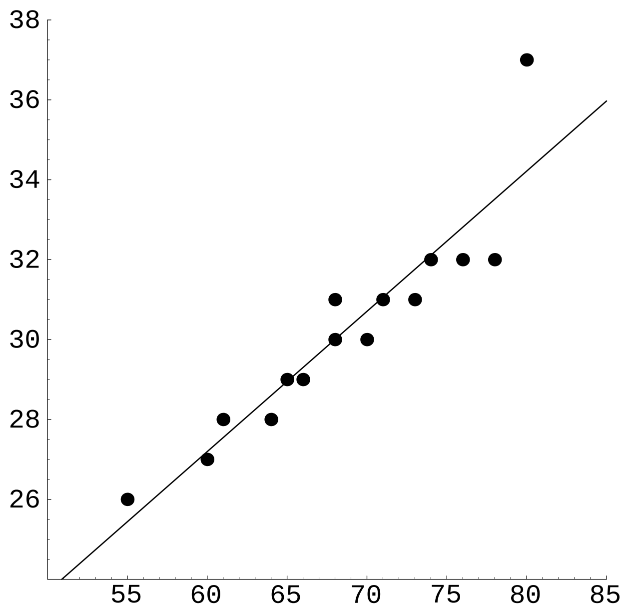

Figura 4.6 Retta di regressione.

Esercizio 4.1 *Calcola la retta di regressione e il coefficiente di Pearson partendo dai dati della Tabella 4.1 ma supponendo che la cavia 3 abbia 55 giorni d'età.*

Osservazione 4.23 Come hai visto (vero?) risolvendo il precedente esercizio, la presenza anche di un solo dato spurio può falsare di molto la retta di regressione, e dare un coefficiente di Pearson molto basso. Per questo motivo nella pratica sperimentale conviene sempre esaminare i dati raccolti per eliminare dati evidentemente spuri, e (possibilmente) investigare i motivi che hanno portato alla presenza dei dati spuri (semplici errori o un fenomeno nuovo?).

4.4 Funzioni polinomiali

Dopo le funzioni quadratiche, si possono considerare funzioni di terzo grado, o di quarto grado, o più in generale *funzioni polinomiali,* cioè funzioni $f\colon \mathbb{R} \to \mathbb{R}$ espresse da un polinomio:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \ , \tag{4.15}$$

dove $n \in \mathbb{N}$ è il *grado* della funzione polinomiale (o del polinomio), e $a_0, \ldots, a_n \in \mathbb{R}$ sono i *coefficienti;* si suppone sempre che $a_n \neq 0$. Non abbiamo ancora gli strumenti necessari[7] per effettuare uno studio dettagliato di queste funzioni; mi limiterò quindi a citare alcuni fatti, in parte analoghi a quanto abbiamo già visto, che possono essere utili nel loro studio. La Figura 4.7 comunque contiene i grafici di alcune funzioni polinomiali, giusto per darti un'idea di che faccia possano avere.

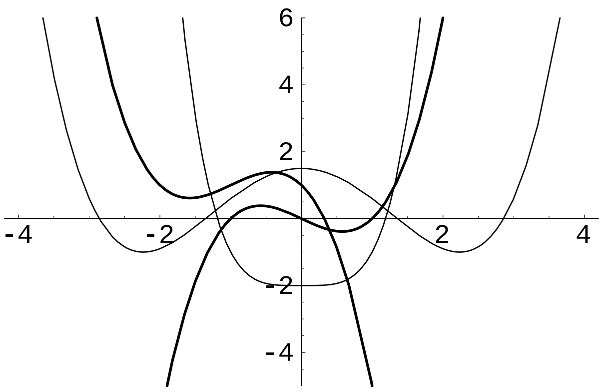

Figura 4.7 Funzioni polinomiali.

La prima osservazione è che per x molto grande in valore assoluto, l'addendo $a_n x^n$ in (4.15) è molto più grande degli altri, per cui il comportamento della funzione f per x molto grande in valore assoluto è dettato dal comportamento di $a_n x^n$. In particolare:

– se $a_n > 0$ e n è pari allora

$$\lim_{x \to \pm\infty} a_n x^n + \cdots + a_0 = \lim_{x \to \pm\infty} a_n x^n = +\infty \ ;$$

[7] Ne introdurremo molti nel prossimo capitolo.

– se $a_n > 0$ e n è dispari allora

$$\lim_{x\to\pm\infty} a_n x^n + \cdots + a_0 = \lim_{x\to\pm\infty} a_n x^n = \pm\infty\ ;$$

– se $a_n < 0$ e n è pari allora

$$\lim_{x\to\pm\infty} a_n x^n + \cdots + a_0 = \lim_{x\to\pm\infty} a_n x^n = -\infty\ ;$$

– se $a_n < 0$ e n è dispari allora

$$\lim_{x\to\pm\infty} a_n x^n + \cdots + a_0 = \lim_{x\to\pm\infty} a_n x^n = \mp\infty\ .$$

Curiosità 4.5 Vediamo come dimostrare correttamente questa affermazione. Poniamo

$$R = \frac{2n \max\{|a_0|, \ldots, |a_n|\}}{|a_n|} \geq 2n > 1\ .$$

Se $|x| > R$ abbiamo $|x|^j > |x| > 2n|a_j|/|a_n|$, cioè

$$\frac{|a_j|}{|a_n|}\,\frac{1}{|x|^j} < \frac{1}{2n}\ ,$$

e quindi

$$\sum_{j=1}^{n} \frac{|a_j|}{|a_n|}\,\frac{1}{|x|^j} < n \cdot \frac{1}{2n} = \frac{1}{2}\ .$$

Ricordando le seguenti fondamentali proprietà del valore assoluto

$$|a| - |b| \leq \Big||a| - |b|\Big| \leq |a + b| \leq |a| + |b|\ , \qquad |a \cdot b| = |a| \cdot |b|\ ,$$

per $|x| > R$ otteniamo

$$1 + \sum_{j=1}^{n} \frac{a_j}{a_n}\,\frac{1}{x^j} \geq 1 - \left|\sum_{j=1}^{n} \frac{a_j}{a_n}\,\frac{1}{x^j}\right| \geq 1 - \sum_{j=1}^{n} \frac{|a_j|}{|a_n|}\,\frac{1}{|x|^j} \geq 1 - \frac{1}{2} = \frac{1}{2}\ .$$

Inoltre,

$$1 + \sum_{j=1}^{n} \frac{a_j}{a_n}\,\frac{1}{x^j} \leq 1 + \left|\sum_{j=1}^{n} \frac{a_j}{a_n}\,\frac{1}{x^j}\right| \leq 1 + \sum_{j=1}^{n} \frac{|a_j|}{|a_n|}\,\frac{1}{|x|^j} \leq 1 + \frac{1}{2} = \frac{3}{2}\ ,$$

sempre per $|x| > R$. Siccome

$$a_n x^n + \cdots + a_0 = a_n x^n \left(1 + \sum_{j=1}^{n} \frac{a_j}{a_n}\,\frac{1}{x^j}\right)\ ,$$

non appena $a_n x^n > 0$ e $|x| > R$ otteniamo

$$\frac{3}{2} a_n x^n \geq a_n x^n + \cdots + a_0 \geq \frac{1}{2}\, a_n x^n \ ,$$

mentre se $a_n x^n < 0$ e $|x| > R$ otteniamo

$$\frac{3}{2} a_n x_n \leq a_n x^n + \cdots + a_0 \leq \frac{1}{2}\, a_n x^n \ .$$

Queste due stime implicano immediatamente le affermazioni volute.

La seconda osservazione riguarda il numero di condizioni necessarie per determinare i coefficienti di un polinomio di grado n. Abbiamo visto che il grafico di una funzione lineare era completamente determinato dal passaggio per due punti, e che il grafico di una funzione quadratica era completamente determinato dal passaggio per tre punti. Analogamente, il grafico di un polinomio di grado n è completamente determinato dal passaggio per $n+1$ punti.

Il metodo per trovare il polinomio dati $n+1$ punti è analogo a quello visto per le funzioni quadratiche. Supponiamo di avere $n+1$ punti $(x_0, y_0), \ldots, (x_n, y_n)$ con ascisse $x_0, \ldots, x_n$ tutte distinte. Trovare un polinomio $f(x) = a_n x^n + \cdots + a_0$ il cui grafico passi per questi punti, cioè tale che $f(x_j) = y_j$ per $j = 0, \ldots, n$, equivale a risolvere il sistema lineare

$$\begin{cases} x_0^n a_n + x_0^{n-1} a_{n-1} + \cdots + a_0 = y_0 \ , \\ \qquad\qquad \vdots \\ x_n^n a_n + x_n^{n-1} a_{n-1} + \cdots + a_0 = y_n \ , \end{cases}$$

di $n+1$ equazioni nelle $n+1$ incognite $a_0, \ldots, a_n$. Per risolvere questo sistema si sottrae ogni equazione dalla successiva; se necessario, si dividono, come in (4.8), i coefficienti del sistema ottenuto per un opportuno fattore comune, utilizzando le formule

$$x^{k+1} - y^{k+1} = (x-y) \sum_{i=0}^{k} x^i y^{k-i} \ ; \tag{4.16}$$

e si ripete il procedimento col nuovo sistema. Dopo n passaggi si arriva a una sola equazione lineare con a_n come unica incognita; ricavata a_n si sostituisce il valore trovato nei sistemi precedenti, ricavando a_{n-1} e poi a_{n-2} e così via fino ad a_0. Questa tecnica per determinare i coefficienti dei polinomi è detta *metodo delle differenze.*

Curiosità 4.6 La formula (4.16) si dimostra col seguente conto:

$$\begin{aligned} (x-y) \sum_{i=0}^{k} x^i y^{k-i} &= \sum_{i=0}^{k} x^{i+1} y^{k-i} - \sum_{i=0}^{k} x^i y^{k-(i-1)} \\ &= x^{k+1} + \sum_{i=0}^{k-1} x^{i+1} y^{k-i} - \sum_{h=0}^{k-1} x^{h+1} y^{k-h} - y^{k+1} \\ &= x^{k+1} - y^{k+1} \ , \end{aligned}$$

dove abbiamo posto $h = i - 1$ nell'ultima sommatoria.

ESEMPIO 4.9 *Studiamo nuovamente i semi di pomodoro degli Esempi 4.1 e 4.7. Sai già che alla temperatura di* 12 °C *germoglia il* 40% *dei semi, alla temperatura di* 15 °C *germoglia il* 70% *dei semi, e che alla temperatura di* 9 °C *germoglia il* 20% *dei semi. Non contento, il tuo assistente effettua un'ulteriore misura, scoprendo che alla temperatura di* 18 °C *germoglia l'*85% *dei semi. Trova un polinomio di terzo grado che rappresenta questi dati.* Dobbiamo trovare a_0, a_1, a_2, $a_3 \in \mathbb{R}$ che soddisfano il sistema

$$\begin{cases} 729a_3 + 81a_2 + 9a_1 + a_0 = 20 \ , \\ 1728a_3 + 144a_2 + 12a_1 + a_0 = 40 \ , \\ 3375a_3 + 225a_2 + 15a_1 + a_0 = 70 \ , \\ 5832a_3 + 324a_2 + 18a_1 + a_0 = 85 \ . \end{cases}$$

La prima serie di sottrazioni ci dà

$$\begin{cases} 999a_3 + 63a_2 + 3a_1 = 20 \ , \\ 1647a_3 + 81a_2 + 3a_1 = 30 \ , \\ 2457a_3 + 99a_2 + 3a_1 = 15 \ . \end{cases}$$

Siccome i coefficienti di a_1 sono uguali nelle tre equazioni, non abbiamo bisogno di effettuare divisioni e possiamo procedere direttamente con la seconda serie di sottrazioni:

$$\begin{cases} 648a_3 + 18a_2 = 10 \ , \\ 810a_3 + 18a_2 = -15 \ . \end{cases}$$

L'ultima sottrazione ci dà $162a_3 = -25$, cioè $a_3 = -25/162$. Mettendo questo valore nelle equazioni precedenti e risalendo troviamo $a_2 = 55/9$, $a_1 = -1265/18$, e $a_0 = 270$, per cui il polinomio cercato è

$$P(T) = -\frac{25}{162}T^3 + \frac{55}{9}T^2 - \frac{1265}{18}T + 270 \ .$$

La Figura 4.8 contiene sia i dati sia i grafici delle funzioni che abbiamo ottenuto nei vari esempi.

Osservazione 4.24 Gli Esempi 4.1, 4.7 e 4.9 mostrano che i conti nel metodo delle differenze sono più semplici se le ascisse $x_0, \ldots, x_n$ sono equispaziate, cioè se $x_1 - x_0 = x_2 - x_1 = \cdots = x_{n-1} - x_n$.

Esercizio 4.2 *Usa il metodo dei minimi quadrati per determinare la retta di regressione per i dati dell'Esempio 4.9, calcola il coefficiente di Pearson, e confronta pregi e difetti delle varie formule (retta di regressione inclusa) che abbiamo trovato per rappresentare la relazione fra temperatura e percentuale di semi germinati.*

L'ultima osservazione che ci servirà riguarda le radici di un polinomio. Una *radice* di un polinomio $f(x)$ è un numero reale $x_0 \in \mathbb{R}$ tale che $f(x_0) = 0$. La regola di Ruffini dice che $x_0 \in \mathbb{R}$ è radice del polinomio f se e solo se esiste un polinomio q

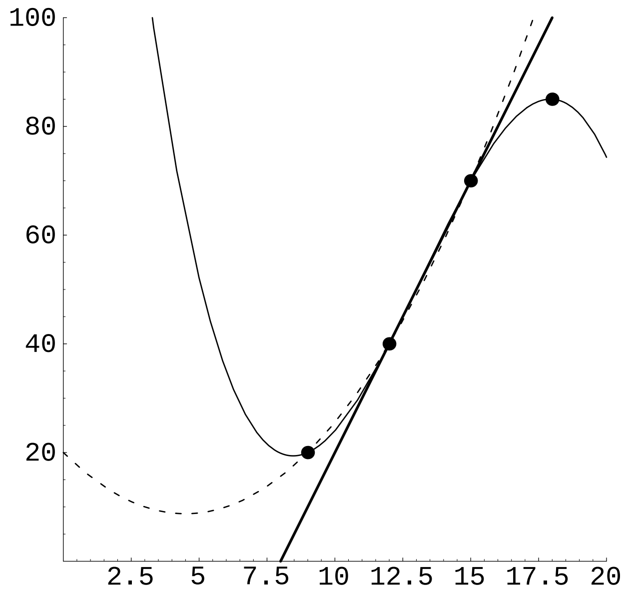

Figura 4.8 .

tale che $f(x) = (x - x_0)q(x)$, dove q ha grado minore di uno rispetto al grado di f. Ora, se x_0 è radice anche di q, deve esistere un terzo polinomio q_1, di grado minore di quello di q, tale che $q(x) = (x - x_0)q_1(x)$, per cui $f(x) = (x - x_0)^2 q_1(x)$. Ripetendo questo procedimento, prima o poi troveremo un numero naturale $r > 0$ (e minore o uguale del grado di f) e un polinomio q_r tali che

$$f(x) = (x - x_0)^r q_r(x) \qquad \text{con} \qquad q_r(x_0) \neq 0 \; ;$$

il numero r è detto *molteplicità* di x_0 come radice di f.

Curiosità 4.7 Se q_r ha una radice x_1 di molteplicità $s > 0$, possiamo ripetere questa costruzione con q_r, trovando un polinomio p_s tale che $f(x) = (x - x_0)^r (x - x_1)^s p_s(x)$. Procedendo in questo modo si riesce a dimostrare che ogni polinomio f si può scrivere in modo unico come prodotto

$$f(x) = a p_1(x)^{r_1} \cdots p_2(x)^{r_k} \; , \tag{4.17}$$

dove $a \in \mathbb{R}$ è il coefficiente direttore di f, $r_1, \ldots, r_k$ sono numeri interi maggiori di zero, e $p_1, \ldots, p_r$ sono polinomi *monici* (cioè con coefficiente direttore uguale a 1) *irriducibili* (cioè non si possono scrivere come prodotto di due altri polinomi monici). I polinomi irriducibili sono l'equivalente per i polinomi dei numeri primi per i numeri naturali; e (4.17) è l'equivalente della decomposizione in fattori primi di un numero intero. Infine, si può anche dimostrare che i polinomi monici irriducibili a coefficienti reali sono o polinomi lineari della forma $x - x_0$ oppure polinomi quadratici senza radici reali (cioè con discriminante negativo).

4.5 Funzioni potenza

Un'altra famiglia importante di funzioni è costituito dalle *funzioni potenza,* che sono funzioni della forma
$$f(x) = ax^p \ ,$$
dove $a \neq 0$ è un numero reale e p è un numero razionale (ma vedi anche l'Osservazione 4.26), detto *esponente* della funzione potenza.

Osservazione 4.25 Se p è un numero naturale, $p \in \mathbb{N} = \{0, 1, 2, \ldots\}$, la funzione $f(x) = ax^p$ è una particolare funzione polinomiale, e quindi è definita su tutta la retta reale: $f: \mathbb{R} \to \mathbb{R}$. Se p è un numero intero negativo, $p \in \mathbb{Z}^- = \{-1, -2, \ldots\}$, la funzione f è una particolare funzione razionale (vedi la prossima sezione) ed è definita per $x \neq 0$, cioè $f: \mathbb{R}^* \to \mathbb{R}$. Infine, se p è un numero razionale non intero, $p \in \mathbb{Q} \setminus \mathbb{Z}$, allora f è definita solo per $x \geq 0$, cioè $f: \mathbb{R}^+ \to \mathbb{R}$.

Osservazione 4.26 Come vedremo nelle Sezioni 4.7 e 4.8 e nel prossimo capitolo, è possibile dare un senso anche alle potenze *irrazionali* di un numero non negativo, per cui potremo considerare funzioni potenza con esponente qualsiasi (ma solo con argomento reale non negativo).

CURIOSITÀ 4.8 Supponiamo che $p \in \mathbb{R} \setminus \mathbb{Q}$ sia un numero irrazionale, e $x \in \mathbb{R}^+$ un numero non negativo. Siccome i numeri razionali possono approssimare bene quanto vogliamo qualsiasi numero reale, un modo per calcolare la potenza irrazionale x^p si basa sul fatto che esiste un numero reale y tale che la potenza *razionale* x^q è arbitrariamente vicina a y non appena q è un numero razionale sufficientemente vicino a p; allora si pone $x^p = y$. Quindi le potenze razionali forniscono approssimazioni arbitrariamente buone delle potenze irrazionali. Un altro modo per esprimere questo concetto è dire che per ogni $\varepsilon > 0$ (arbitrariamente piccolo) esiste un $\delta > 0$ (sufficientemente piccolo) tale che se q è un numero razionale che dista da p meno di δ (cioè $|q - p| < \delta$) allora x^q dista da y meno di ε (cioè $|x^q - y| < \varepsilon$).

Le funzioni potenza ax^p con $p \in \mathbb{N}$ hanno un comportamento molto simile a quello di ax se p è dispari, e a quello di ax^2 se p è pari. Infatti, con le tecniche viste nelle Sezioni 4.1 e 4.2 è facile dimostrare (esercizio per te e il tuo assistente) che

– se p è dispari, la funzione $f(x) = ax^p$ è monotona (crescente se $a > 0$, decrescente se $a < 0$), e
$$\lim_{x \to \pm\infty} ax^p = \begin{cases} \pm\infty & \text{se } a > 0 \ , \\ \mp\infty & \text{se } a < 0 \ ; \end{cases}$$
– se p è pari (e non nullo), la funzione $f(x) = ax^p$ ha un punto di minimo (se $a > 0$) o un punto di massimo (se $a < 0$) in $x = 0$, è monotona (crescente o decrescente a seconda del segno di a), nelle semirette $(-\infty, 0]$ e $[0, +\infty)$, e
$$\lim_{x \to \pm\infty} ax^p = \begin{cases} +\infty & \text{se } a > 0 \ , \\ -\infty & \text{se } a < 0 \ . \end{cases}$$

Osservazione 4.27 Nota che il tipo di monotonia su $\mathbb{R}^+ = [0, +\infty)$ e il limite a $+\infty$ di ax^p dipende solo dal segno di a e non dalla parità di p.

Osservazione 4.28 Il confronto fra funzioni potenza con diverso esponente può fornire interessanti conseguenze biologiche legate a questioni di scala. Prendiamo un particolare individuo di una data specie animale (o vegetale), e scegliamo una sua lunghezza caratteristica ℓ: per esempio, possiamo indicare con ℓ il diametro della testa. Tutte le altre lunghezze di questo particolare individuo sono banalmente proporzionali a ℓ: i coefficienti di proporzionalità si ottengono semplicemente dividendo per ℓ la lunghezza che si vuole considerare. Se ora prendiamo un altro individuo della stessa specie, i coefficienti di proporzionalità delle sue lunghezze rispetto a ℓ saranno un po' diversi, ma non *troppo* diversi: per esempio, il secondo individuo sarà un po' più alto o un po' più basso, ma facendo parte della stessa specie è molto improbabile che sia alto più del doppio o meno della metà. Quindi possiamo considerare questa lunghezza ℓ come *rappresentativa* della specie; tutte le altre lunghezze in tutti gli altri individui della stessa specie saranno proporzionali a ℓ con coefficienti di proporzionalità approssimativamente costanti. Di conseguenza, tutte le *superfici* di individui della stessa specie saranno proporzionali a ℓ^2, con coefficienti di proporzionalità approssimativamente costanti; e tutti i *volumi* saranno proporzionali a ℓ^3, con coefficienti di proporzionalità approssimativamente costanti. Ora, i fenomeni di scambio con l'esterno (assorbimento di ossigeno, emissione di calore, eccetera) di un individuo avvengono usualmente attraverso la superficie, e quindi avranno andamenti proporzionali a ℓ^2; invece, i fenomeni metabolici (consumo di ossigeno, produzione di calore, eccetera) sono di solito proporzionali al volume (al numero di cellule coinvolte), e quindi proporzionali a ℓ^3. La conseguenza di tutto ciò è che (come sarà chiarito dai prossimi esempi) *non è possibile variare eccessivamente le dimensioni di una data specie animale senza danneggiare l'equilibrio fra il metabolismo interno e l'ambiente esterno che le permette di vivere;* in un certo senso, i rapporti fra il metabolismo e la forma di una specie ne determinano le dimensioni ideali.

Esempio 4.10 Con buona pace dei film dell'orrore, un ragno gigante, ottenuto ingrandendo 100 volte un ragno usuale, ha poche possibilità di sopravvivere. Il consumo di ossigeno è proporzionale al volume del ragno, volume che è passato da un multiplo di ℓ^3 a un uguale multiplo di $(10\ell)^3 = 1000\ell^3$. D'altra parte, l'assorbimento di ossigeno è proporzionale alla superficie interna dei polmoni, superficie che è passata da un multiplo di ℓ^2 a un uguale multiplo di $(10\ell)^2 = 100\ell^2$. Il consumo di ossigeno è quindi aumentato di 1000 volte, mentre l'assorbimento di ossigeno solo di 100 volte; il ragno gigante riceve solo 1/10 dell'ossigeno che gli servirebbe, e quindi muore soffocato sotto gli occhi increduli dell'eroe del film.

In maniera analoga, con buona pace dei film di fantascienza, un uomo rimpicciolito di 10 volte si sentirebbe piuttosto male. Infatti, la perdita di calore attraverso l'epidermide è proporzionale alla superficie del corpo, cioè a ℓ^2, e quindi si è ridotta di 1/100. Ma il calore prodotto dal corpo umano, necessario alla nostra sopravvivenza visto che siamo animali a sangue caldo, è proporzionale al volume del corpo, cioè a ℓ^3, e quindi si è ridotto di 1/1000. Quindi un uomo rimpicciolito 10 volte perderebbe attraverso l'epidermide 10 volte più calore di quello che produce, e quindi probabilmente morirebbe di freddo.

Osservazione 4.29 La superficie di assorbimento dell'ossigeno nei polmoni in realtà ha una struttura frastagliata, di tipo frattale, e di conseguenza che l'assorbimento di ossigeno è più efficiente diventando proporzionale a ℓ^p con $p > 2$; ma in ogni caso non raggiunge ℓ^3.

ESEMPIO 4.11 È noto che la forza muscolare di un muscolo delle gambe è approssimativamente proporzionale al numero di fibre muscolari che lo compongono, e quindi è proporzionale alla superficie trasversa del muscolo, cioè a ℓ^2. Inoltre, l'energia prodotta dal muscolo è proporzionale alla forza per la lunghezza, e quindi è uguale a $c_1\ell^3$ per un'opportuna costante $c_1 > 0$. D'altra parte, l'energia necessaria per un salto di altezza h è proporzionale al prodotto dell'altezza per il peso del corpo, e quindi è uguale a $c_2h\ell^3$ per un'opportuna costante $c_2 > 0$. Ne segue che la massima altezza possibile h di un salto deve soddisfare $c_1\ell^3 = c_2h\ell^3$, cioè $h = c_1/c_2$; in particolare, h non dipende da ℓ.

Ora, una pulce comune è in grado di saltare a un'altezza pari a 200 volte la propria altezza; quanto sarà in grado di saltare una pulce 10 volte più grande? E una pulce 100 volte più grande? La pulce comune è in grado di saltare a un'altezza $h = 200c_3\ell$ per una costante $c_3 > 0$ opportuna. Una pulce 10 volte più grande avrà lunghezza caratteristica 10ℓ, ma uguali costanti di proporzionalità c_1, c_2 e c_3. Siccome la massima altezza possibile per un salto dipende solo dalle costanti di proporzionalità, la pulce 10 volte più grande *può saltare solo alla stessa altezza della pulce comune,* e quindi al massimo a 20 volte la propria altezza. Analogamente, una pulce 100 volte più grande riuscirà a saltare solo il doppio della propria altezza, e una pulce mastodontica 1000 volte più grande della pulce comune riuscirà a saltare solo un quinto della propria altezza. Per intenderci, se una pulce comune alta circa 1 mm riesce a saltare 20 cm, una pulce alta 1 metro riuscirebbe a saltare sempre soltanto 20 cm…

Le funzioni $f(x) = ax^p$ con esponente $p \in \mathbb{Q}^+ \setminus \mathbb{N}$ razionale positivo non intero sono definite solo su $\mathbb{R}^+$, e hanno comportamento analogo a quelle con esponente naturale: sono crescenti se $a > 0$, decrescenti se $a < 0$, e il loro limite all'infinito è uguale a $\pm\infty$ a seconda del segno di a. La Figura 4.9 contiene il grafico di alcune di queste (con $a = 1$). Nota che quelle con esponente maggiore di 1 hanno la concavità rivolta verso l'alto, mentre quelle con esponente minore di 1 hanno la concavità rivolta verso il basso. Nel prossimo capitolo vedremo come verificare rigorosamente questa affermazione.

Più interessanti sono le funzioni $f(x) = ax^p$ con $p \in \mathbb{Q}^-$; siccome (almeno quando $p \in \mathbb{Z}^-$) sono funzioni razionali, le discutiamo nella prossima sezione (Osservazione 4.31).

4.6 Funzioni razionali

Una *funzione razionale* è un quoziente di polinomi:

$$f(x) = \frac{a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0}{b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0} \ ,$$

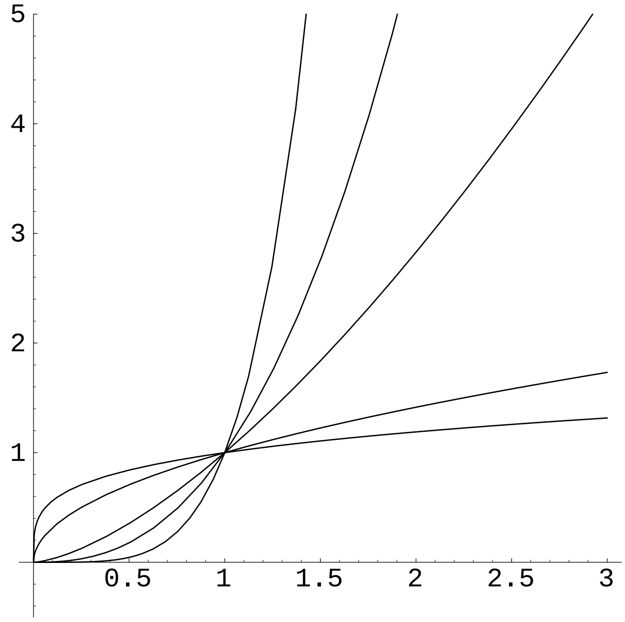

Figura 4.9 Funzioni potenza.

con m, $n \in \mathbb{N}$, $a_0, \ldots, a_m$, $b_0, \ldots, b_n \in \mathbb{R}$ e a_m, $b_n \neq 0$ (e di solito si assume anche $n \geq 1$, perché altrimenti f sarebbe un polinomio); il numero $d = \max\{m, n\}$ è detto *grado* della funzione razionale. Ovviamente, lo studio dell'andamento delle funzioni razionali generiche, come per i polinomi, richiede strumenti che ancora non abbiamo; a parte alcune osservazioni finali, ci concentreremo quindi sulle funzioni razionali di grado 1, note anche come *funzioni lineari fratte.*

L'esempio più semplice di funzione lineare fratta è la funzione (potenza)

$$f(x) = \frac{a}{x} = ax^{-1} ,$$

con $a \neq 0$. Rappresenta le relazioni di proporzionalità inversa: infatti, un punto (x, y) appartiene al grafico di f se e solo se

$$xy = a ,$$

per cui il prodotto fra l'argomento e il valore della funzione è costante su tutto il dominio della funzione.

Osservazione 4.30 In particolare, basta conoscere un punto (x_0, y_0) del grafico per determinare la funzione a/x, in quanto $a = x_0 y_0$.

La prima osservazione importante è che la funzione $f(x) = a/x$ *non è definita su tutto l'asse reale:* il suo dominio *non* è $\mathbb{R}$. Infatti, il quoziente a/x non è definito per $x = 0$, per cui la funzione f è definita solo su $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$. Si dice anche che 0 è una *singolarità* per la funzione f.

Quando una funzione ha una singolarità (un punto in cui non è definita), è importante cercare di capire come si comporta vicino alla singolarità. Cominciamo supponendo $a > 0$ e $x > 0$. Quando $x > 0$ diventa piccolo (per esempio, minore di $1/n$ per n arbitrariamente grande) allora $1/x$ diventa grande (per esempio, maggiore di n) e quindi anche a/x diventa grande (per esempio, maggiore di an). In altre parole, *possiamo rendere $f(x) = a/x$ arbitrariamente grande a patto di scegliere x sufficientemente piccolo e positivo.*

Abbiamo già visto come tradurre in simboli i concetti di "arbitrariamente grande" e "sufficientemente grande"; una procedura analoga si usa per il concetto di "sufficientemente piccolo". La frase precedente diventa: per ogni $M > 0$ (arbitrariamente grande) esiste $\delta > 0$ (sufficientemente piccolo) tale che se $0 < x < \delta$ allora $f(x) > M$. In simboli,

$$\forall M > 0 \quad \exists \delta > 0\colon \quad 0 < x < \delta \Longrightarrow f(x) > M \ .$$

Usando la terminologia dei limiti, diremo che *il limite di $f(x)$ per x che tende a* 0 *da destra* (o *da sopra,* o *che tende a* 0^+) *è* $+\infty$*:*

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

Quando $x < 0$ negativo diventa piccolo (in valore assoluto), $1/x$ diventa grande in valore assoluto ma rimane negativo, cioè diventa molto negativo. Quindi, sempre assumendo $a > 0$, *possiamo rendere $f(x) = a/x$ arbitrariamente negativa a patto di scegliere x sufficientemente piccolo e negativo.* In altre parole, per ogni $M > 0$ (arbitrariamente grande) esiste $\delta > 0$ (sufficientemente piccolo) tale che se $-\delta < x < 0$ allora $f(x) < -M$. In simboli,

$$\forall M > 0 \quad \exists \delta > 0\colon \quad -\delta < x < 0 \Longrightarrow f(x) < -M \ .$$

Usando la terminologia dei limiti, diremo che *il limite di $f(x)$ per x che tende a* 0 *da sinistra* (o *da sotto,* o *che tende a* 0^-) *è* $-\infty$*:*

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

Una conseguenza di questo comportamento è che il grafico di f si avvicina sempre più all'asse delle ordinate quando x tende a zero (si avvicina in alto se $x > 0$, in basso se $x < 0$). Si dice che l'asse delle ordinate è un *asintoto verticale* della funzione f.

Ovviamente, se $a < 0$ i segni si invertono; lascio a te il compito di dimostrare che

$$\text{se } a < 0 \text{ allora } \lim_{x \to 0^\pm} f(x) = \mp\infty \ ,$$

e di scrivere a parole e in simboli il significato di questa formula.

Vediamo ora cosa succede allontanandoci da zero. Cominciamo come al solito supponendo $a > 0$. Se $0 < x_0 < x_1$ allora $0 < 1/x_1 < 1/x_0$ e $0 < a/x_1 < a/x_0$; quindi $0 < x_0 < x_1$ implica $f(x_0) > f(x_1) > 0$, per cui f è strettamente decrescente nella semiretta $(0, +\infty)$. In modo analogo si dimostra che f è strettamente decrescente (ma negativa) nella semiretta $(-\infty, 0)$. Se invece $a < 0$ s'inverte tutto; riassumendo,

– se $a > 0$, la funzione $f(x) = a/x$ è strettamente decrescente e negativa in $(-\infty, 0)$, mentre è strettamente decrescente e positiva in $(0, +\infty)$;
– se $a < 0$, la funzione $f(x) = a/x$ è strettamente crescente e positiva in $(-\infty, 0)$, mentre è strettamente crescente e negativa in $(0, +\infty)$.

In particolare, quando x diventa grande (e $a > 0$) la funzione $f(x) = a/x$ decresce rimanendo positiva e diventando arbitrariamente piccola. Una cosa simile accade quando x diventa molto negativo (e quando $a < 0$): cambia il segno, cambia la crescenza, ma in ogni caso $f(x)$ *diventa arbitrariamente piccolo in valore assoluto a patto di scegliere x sufficientemente grande o sufficientemente negativo.* Ormai avrai capito il trucco per tradurre espressioni quali "arbitrariamente piccolo": la frase precedente diventa "per ogni $\varepsilon > 0$ (arbitrariamente piccolo) esiste $M > 0$ (sufficientemente grande) tale che se $x > M$ o $x < -M$ allora $|f(x)| < \varepsilon$". In simboli,

$$\forall\varepsilon > 0 \quad \exists M > 0\colon \quad x > M \text{ o } x < -M \implies |f(x)| < \varepsilon\ ,$$

o anche

$$\lim_{x\to\pm\infty} f(x) = 0\ ,$$

che si legge "il limite di $f(x)$ per x che tende a $\pm\infty$ è 0", o anche "$f(x)$ tende a 0 per x che tende a $\pm\infty$". Una conseguenza di questo comportamento è che il grafico di f si avvicina sempre più all'asse delle ascisse quando x tende a $\pm\infty$ (si avvicina da sopra a $+\infty$ se $a > 0$ e a $-\infty$ se $a < 0$, e da sotto a $-\infty$ se $a > 0$ e a $+\infty$ se $a < 0$). Si dice che l'asse delle ascisse è un *asintoto orizzontale* della funzione f. La Figura 4.10 contiene il grafico di $f(x) = a/x$ con $a > 0$, che riassume visivamente tutte le proprietà che abbiamo discusso, asintoti compresi. Questo grafico è un esempio di *iperbole equilatera*[8].

Osservazione 4.31 Le funzioni potenza $f(x) = ax^p$ con p razionale negativo possono venire studiate con tecniche analoghe. In particolare, su $(0, +\infty)$ sono tutte positive e strettamente decrescenti (se $a > 0$, o negative e crescenti se $a < 0$), hanno limite 0 a $+\infty$ con la retta delle ascisse come asintoto orizzontale, e limite $\pm\infty$ (a seconda del segno di a) per x che tende a 0^+, con la retta delle ordinate come asintoto verticale. Su $(-\infty, 0)$ sono definite solo quando $p \in \mathbb{Z}^-$, e in tal caso hanno un andamento analogo a quello di a/x in $(-\infty, 0)$ se p è dispari, e a quello

[8] Il termine "equilatera" serve a indicare che i due asintoti sono ortogonali. In generale, un'*iperbole* è il luogo dei punti del piano per cui il valore assoluto delle differenze delle distanze da due punti dati (detti *fuochi*) è costante. Un'iperbole ha sempre due asintoti, ma non sono necessariamente ortogonali.

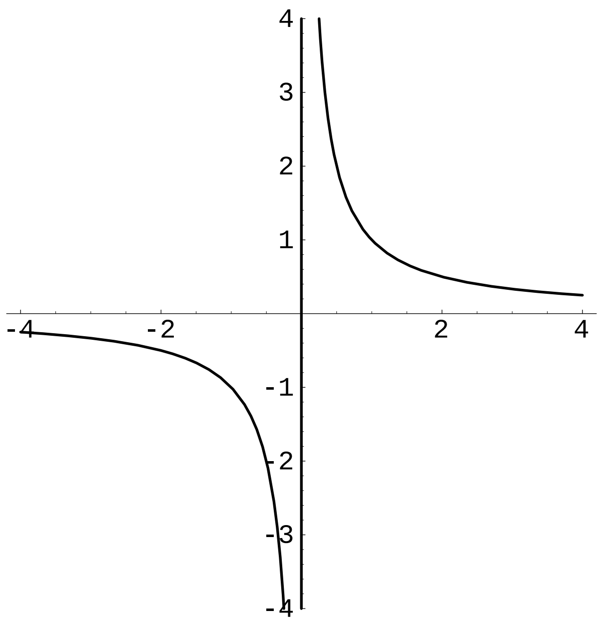

Figura 4.10 $f(x) = 1/x$.

di $|a/x|$ se p è pari. In particolare, se p è pari abbiamo che $f(x)$ diventa arbitrariamente grande a patto di prendere x sufficientemente piccolo in valore assoluto, cioè per ogni $M > 0$ (arbitrariamente grande) esiste $\delta > 0$ (sufficientemente piccolo) tale che $0 < |x| < \delta$ implica $f(x) > M$, ovvero

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

senza bisogno di distinguere se x tende a 0 da destra o da sinistra.

Lo studio di qualsiasi funzione razionale fratta può venire ricondotto a quello di a/x, proprio come avevamo ricondotto lo studio di qualsiasi funzione quadratica a quello di x^2. Infatti, se $c \neq 0$ si ha

$$\frac{ax+b}{cx+d} = \frac{(a/c)x + b/c}{x + d/c} = \frac{\frac{a}{c}(x + d/c) + \frac{b}{c} - \frac{ad}{c^2}}{x + d/c} = \frac{a}{c} + \frac{(bc - ad)/c^2}{x + d/c} \,. \qquad (4.18)$$

Ricordando la Sezione 4.2 vediamo che il grafico di $f(x) = (ax + b)/(cx + d)$ si ottiene a partire da quello di $1/x$ con le seguenti operazioni:

- moltiplichiamo le ordinate per $(bc - ad)/c^2$;
- sottraiamo a/c alle ordinate, traslando il grafico in direzione verticale della quantità a/c;
- sommiamo d/c alle ascisse, traslando il grafico in direzione orizzontale della quantità $-d/c$.

Quindi la funzione $f(x) = (ax + b)/(cx + d)$ soddisfa le seguenti proprietà:

- ha una singolarità in $x_0 = -d/c$;
- il suo grafico è un'iperbole equilatera con asintoto orizzontale la retta $y = a/c$ e asintoto verticale la retta $x = -d/c$;
- nelle semirette $(-\infty, -d/c)$ e $(-d/c, +\infty)$ è strettamente decrescente se si ha $bc - ad > 0$, mentre se $bc - ad < 0$ è strettamente crescente nelle stesse semirette (ed è costante se $bc - ad = 0$).

La presenza dell'asintoto orizzontale $y = a/c$ vuol dire che $f(x)$ è arbitrariamente vicina al valore a/c non appena x è sufficientemente grande o sufficientemente negativo. Ora, $f(x)$ è arbitrariamente vicina ad a/c se e solo se la differenza $f(x) - a/c$ è arbitrariamente piccola in valore assoluto. Quindi dire che $y = a/c$ è un asintoto orizzontale per f equivale a dire che *per ogni* $\varepsilon > 0$ *(arbitrariamente piccolo) esiste* $M > 0$ *(sufficientemente grande) tale che se* $x > M$ *o* $x < -M$ *allora* $|f(x) - a/c| < \varepsilon$. In simboli,

$$\forall \varepsilon > 0 \quad \exists M > 0\colon \quad x > M \text{ o } x < -M \Longrightarrow |f(x) - a/c| < \varepsilon \,,$$

o anche

$$\lim_{x \to \pm\infty} f(x) = a/c \,.$$

Analogamente, la presenza dell'asintoto verticale $x = -d/c$ vuol dire che $f(x)$ è arbitrariamente grande (o arbitrariamente negativa) non appena x è sufficientemente vicino a $x_0 = -d/c$. Ora, x è sufficientemente vicina a x_0 se e solo se la differenza $x - x_0$ è arbitrariamente piccola. Quindi dire che $x = x_0$ è un asintoto verticale per f equivale (almeno quando $bc - ad > 0$) a dire che *per ogni* $M > 0$ *(arbitrariamente grande) esiste* $\varepsilon > 0$ *(sufficientemente piccolo) tale che se* $0 < x - x_0 < \varepsilon$ *allora* $f(x) > M$*, e se* $-\varepsilon < x - x_0 < 0$ *allora* $f(x) < -M$. Usando il simbolo di limite questo si scrive

$$\lim_{x \to x_0^+} f(x) = +\infty \quad \text{e} \quad \lim_{x \to x_0^-} f(x) = -\infty \,.$$

Lascio a te il compito di scrivere cosa succede se $bc - ad < 0$. La Figura 4.11 contiene il grafico di una funzione lineare fratta con rappresentati anche gli asintoti.

Osservazione 4.32 Una conseguenza immediata di (4.18) è che i punti (x, y) del grafico di $f(x) = (ax + b)/(cx + d)$ sono tutti i punti del piano che soddifano la condizione

$$(x - \alpha)(y - \beta) = k$$

con $\alpha = -d/c$, $\beta = a/c$ e $k = (bc - ad)/c^2$.

Osservazione 4.33 Abbiamo visto cosa vuol dire che una funzione ha limite infinito quando x tende all'infinito; cosa vuol dire che ha limite un valore finito quando x tende all'infinito; e cosa vuol dire che ha limite infinito quando x tende a un valore finito (da destra, da sinistra o da entrambi i lati). Rimane da dire cosa vuol dire che ha limite un valore finito quando x tende a un valore finito. La definizione non

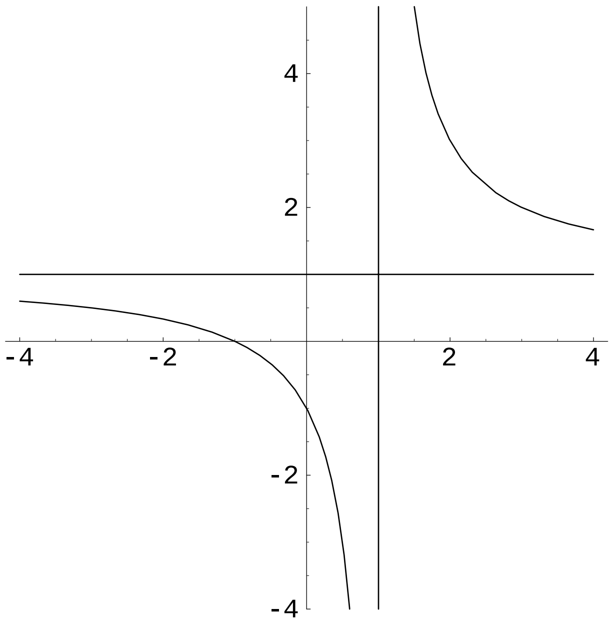

Figura 4.11 $f(x) = (x+1)/(x-1)$.

dovrebbe stupirti: diremo che la funzione f ha *limite* ℓ *quando* x *tende a* x_0, e scriviamo

$$\lim_{x \to x_0} f(x) = \ell$$

se $f(x)$ si avvicina arbitrariamente a ℓ a patto di prendere x sufficientemente vicino a x_0, o, in altre parole, se *per ogni* $\varepsilon > 0$ *(arbitrariamente piccolo) esiste* $\delta > 0$ *(sufficientemente piccolo) tale che* $0 < |x - x_0| < \delta$ *implica* $|f(x) - \ell| < \varepsilon$.

Osservazione 4.34 La condizione $0 < |x - x_0| < \delta$ è equivalente a richiedere $0 < x - x_0 < \delta$ oppure $-\delta < x - x_0 < 0$. Utilizzando solo una di queste due condizioni otteniamo (come visto prima) il concetto di limite per x che tende a x_0 da sopra (cioè $x \to x_0^+$) oppure da sotto (cioè $x \to x_0^-$).

CURIOSITÀ 4.9 Possiamo usare il concetto di limite per definire la continuità di una funzione. Per l'esattezza, una funzione $f\colon I \to \mathbb{R}$, dove $I \subseteq \mathbb{R}$ è un intervallo, è *continua* in un punto $x_0 \in I$ se

$$\lim_{x \to x_0} f(x) = f(x_0) ,$$

cioè se il suo valore in x_0 coincide con il suo limite (sia da sopra che da sotto) in x_0, o, ancora, se $f(x)$ diventa arbitrariamente vicino a $f(x_0)$ a patto di prendere x sufficientemente vicino a x_0. La funzione $f\colon I \to \mathbb{R}$ è poi detta *continua* se lo è in ogni punto del suo dominio I.

Vediamo ora un'applicazione biologica delle funzioni lineari fratte.

Esempio 4.12 In un esperimento si trova[9] che la velocità v (in cm/sec) con cui un muscolo sartorio della coscia di una rana si espande per sollevare un peso p (in grammi) soddisfa la relazione

$$v(p) = 0.95\left(\frac{70-p}{p+12}\right) .$$

In particolare, questa funzione ha una singolarità in $p = -12$; ma siccome chiaramente ci interessa solo per $p \geq 0$, la presenza della singolarità non è un problema. È una funzione lineare fratta della forma $(ax+b)/(cx+d)$ con $a = -0.95$, $b = 66.5$, $c = 1$ e $d = 12$. In particolare, $(bc-ad)/c^2 = 77.9 > 0$ per cui v è strettamente decrescente per $p > -12$ (in altre parole, maggiore il peso più lentamente si estende il muscolo, osservazione piuttosto ragionevole). Di conseguenza, la massima velocità di estensione si ha per $p = 0$, cioè in assenza di carico, e vale $v(0) \simeq 5.54$ cm/sec. Il limite all'infinito $a/c = -0.95$ è negativo; dunque il grafico deve intersecare l'asse delle ascisse. Infatti $v(70) = 0$, che vuol dire che se $p = 70$ g la gamba della rana non riesce a espandersi (velocità zero!), cioè la rana non riesce a sollevare un peso di 70 g (o maggiore). In particolare, questa formula può essere valida solo per valori di p nell'intervallo $[0, 70]$.

Vediamo ora cosa possiamo dire sul comportamento di una funzione razionale qualsiasi

$$f(x) = \frac{p(x)}{q(x)} = \frac{a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0}{b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0} . \tag{4.19}$$

Cominciamo col capire dove è definita. Gli unici punti in cui potrebbero esserci dei problemi sono le radici del denominatore q. Se x_0 è una radice di q e il numeratore non si annulla in x_0, cioè $p(x_0) \neq 0$, allora per calcolare $f(x_0)$ dovremmo dividere per zero, che non è possibile; quindi *le radici del denominatore che non sono radici del numeratore sono sicuramente singolarità* per f. Se invece x_0 è radice anche del numeratore, dobbiamo confrontare le molteplicità. Nella Sezione 4.4 abbiamo visto che se x_0 è radice sia di p che di q possiamo scrivere $p(x) = (x-x_0)^r p_1(x)$ e $q(x) = (x-x_0)^s q_1(x)$ con $p_1(x_0)$, $q_1(x_0) \neq 0$. Quindi

$$f(x) = \frac{(x-x_0)^r p_1(x)}{(x-x_0)^s q_1(x)} = (x-x_0)^{r-s}\,\frac{p_1(x)}{q_1(x)} . \tag{4.20}$$

Ma allora se $r \geq s$ il punto x_0 non è una singolarità di f, in quanto $q_1(x_0) \neq 0$; invece x_0 rimane una singolarità di f se $r < s$, perché in tal caso è $(x-x_0)^{r-s}$ ad avere una singolarità in x_0 (l'esponente è negativo). Inoltre (4.20) implica che, semplificando un'opportuna potenza di $x - x_0$, possiamo esprimere f come quoziente di polinomi tali che x_0 non sia una radice comune di numeratore e denominatore. Ripetendo questo procedimento per tutte le radici del denominatore troviamo che

[9] Nella Sezione 4.9 vedremo come il metodo dei minimi quadrati può essere usato anche per interpolare funzioni lineari fratte.

ogni funzione razionale si può esprimere come rapporto di polinomi privi di radici comuni, e in tal caso le singolarità coincidono con le radici del denominatore.

Supponiamo allora che $f(x) = p(x)/q(x)$ sia una funzione razionale tale che p e q non abbiano radici comuni, e sia x_0 una radice del denominatore di molteplicità $r > 0$. Per quanto visto possiamo scrivere

$$f(x) = \frac{1}{(x-x_0)^r}\,\frac{p(x)}{q_1(x)} \qquad \text{con} \qquad p(x_0),\, q_1(x_0) \neq 0\;.$$

In particolare, $p(x)/q_1(x)$ ammette limite finito non nullo $\ell = p(x_0)/q_1(x_0)$ per x che tende a x_0, per cui il comportamento di $f(x)$ per x vicino a x_0 sarà analogo a quello di $\ell/(x-x_0)^r$. In particolare, il limite di $f(x)$ per x che tende a x_0 (da sopra o da sotto) sarà uguale a quello di $\ell/(x-x_0)^r$, e quindi varrà $\pm\infty$ a seconda del segno di ℓ e della parità di r (e se x tende a x_0 da sopra o da sotto). In ogni caso, la retta $x = x_0$ è un asintoto verticale.

Osservazione 4.35 Nel ragionamento precedente abbiamo implicitamente usato alcune proprietà algebriche dei limiti. Per l'esattezza, le seguenti formule valgono *quasi* sempre:

$$\begin{aligned} \lim_{x\to x_0}[f(x)+g(x)] &= \lim_{x\to x_0} f(x) + \lim_{x\to x_0} g(x)\;, \\ \lim_{x\to x_0}[f(x)g(x)] &= \lim_{x\to x_0} f(x)\cdot \lim_{x\to x_0} g(x)\;, \\ \lim_{x\to x_0}\frac{f(x)}{g(x)} &= \frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)} \qquad \text{se } \lim_{x\to x_0} g(x) \neq 0\;. \end{aligned} \tag{4.21}$$

Il "quasi" si riferisce al fatto che possono sorgere dei problemi se qualcuno di questi limiti è infinito. In particolare, se a secondo membro otteniamo una delle seguenti *forme indeterminate*

$$+\infty - \infty\;, \qquad \pm\infty\cdot 0\;, \qquad \frac{\pm\infty}{\pm\infty}\;,$$

allora per scoprire quanto fa il limite a primo membro (ammesso che esista!) abbiamo bisogno di maggiori informazioni (e di metodi che vedremo nel prossimo capitolo). Altre apparizioni di limiti infiniti non creano grossi problemi, invece: se poniamo

$$\begin{gathered} +\infty + \infty = +\infty\;, \quad -\infty - \infty = -\infty\;, \\ \ell \pm \infty = \pm\infty \quad \text{per } \ell \in \mathbb{R}\;, \\ \ell\cdot\pm\infty = \frac{\pm\infty}{\ell} = \pm\infty \quad \text{per } \ell > 0\;, \qquad \ell\cdot\pm\infty = \frac{\pm\infty}{\ell} = \mp\infty \quad \text{per } \ell < 0\;, \end{gathered}$$

allora le formule (4.21) rimangono valide (come pure rimangono valide se al posto di x_0 mettiamo $\pm\infty$). Infine, anche lo studio del limite di un quoziente quando il denominatore tende a zero richiede maggiori informazioni. Se il limite del numeratore è non nullo (o infinito), il limite del *valore assoluto* del quoziente è $+\infty$, ma il limite del quoziente potrebbe essere $+\infty$, $-\infty$ o non esistere affatto. Infine, se anche il limite del numeratore è nullo siamo in presenza della forma indeterminata $0/0$, che studieremo nel prossimo capitolo.

Infine, vediamo cosa possiamo dire sul comportamento di una funzione razionale f, scritta nella forma (4.19), quando x tende all'infinito. Abbiamo visto (nella Sezione 4.4) che per $|x|$ abbastanza grande, $p(x)$ si comporta come $a_m x^m$ e $q(x)$ si comporta come $b_n x^n$; quindi $f(x)$ si comporta come

$$\frac{a_m x^m}{b_n x^n} = \frac{a_m}{b_n} x^{m-n}$$

quando $|x|$ è abbastanza grande.

Curiosità 4.10 Per l'esattezza, quanto visto nella Curiosità 4.5 ci dice che esiste un $R > 0$ sufficientemente grande tale che se $|x| > R$ e $a_m x^m, b_n x^n > 0$ allora $\frac{1}{2} a_m x^m \leq p(x) \leq \frac{3}{2} a_m x^m$ e $\frac{1}{2} b_n x^n \leq q(x) \leq \frac{3}{2} b_n x^n$, per cui

$$\frac{1}{3} \frac{a_m}{b_n} x^{m-n} \leq \frac{p(x)}{q(x)} = f(x) \leq 3 \frac{a_m}{b_n} x^{m-n} .$$

Stime analoghe si ottengono anche per gli altri possibili segni di $a_m x^m$ e $b_n x^n$.

Di conseguenza,

$$\lim_{x \to \pm\infty} f(x) = \lim_{x \to \pm\infty} \frac{a_m}{b_n} x^{m-n} = \begin{cases} 0 & \text{se } n > m , \\ \dfrac{a_m}{b_n} & \text{se } n = m , \\ \pm\infty & \text{se } n < m , \end{cases}$$

dove il segno nell'ultimo caso dipende dal segno di a_m/b_n, dalla parità di $m-n$, e da dove si sta calcolando il limite (se a $+\infty$ o a $-\infty$).

Esempio 4.13 La legge che descrive il comportamento delle lenti convesse sottili è

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f} , \tag{4.22}$$

dove u è la distanza fra l'oggetto e il centro della lente, v è la distanza fra l'immagine e il centro della lente, e f è la lunghezza focale. Vogliamo studiare la dipendenza della distanza $s = u + v$ fra oggetto e immagine dalla distanza u dell'oggetto da una lente di lunghezza focale $f = 10$ cm. Siccome $v = s - u$, otteniamo

$$\frac{1}{10} = \frac{1}{u} + \frac{1}{s-u} = \frac{s}{u(s-u)} \quad \Longrightarrow \quad u(s-u) = 10s \quad \Longrightarrow \quad s = \frac{u^2}{u-10} .$$

Quindi la dipendenza è data da una funzione razionale di grado 2. Notiamo prima di tutto che s ha una singolarità in $u = 10$: questo vuol dire che quando l'oggetto si avvicina al fuoco della lente allora l'immagine scappa all'infinito (in quanto $s(u)$ tende a $+\infty$ per u che tende a 10^+). Se $u < 10$ allora s diventa negativa, cosa fisicamente insensata; ma infatti la legge (4.22) vale solo per u, $v \geq f$.

Siccome il grado del numeratore è maggiore del grado del denominatore, e i coefficienti dei termini di grado massimo del numeratore e del denominatore sono positivi, abbiamo che s tende a $+\infty$ quando u tende all'infinito. Detta così è un'affermazione fisicamente ovvia: se l'oggetto scappa all'infinito, la distanza dall'immagine (che è dall'altro lato della lente) tende all'infinito. Possiamo renderla però più interessante notando che

$$\frac{u^2}{u-10} = \frac{u(u-10)+10u}{u-10} = u + \frac{10(u-10)+100}{u-10} = u + 10 + \frac{100}{u-10} .$$

Siccome $100/(u-10)$ tende a zero quando u tende a $+\infty$, vediamo che s si comporta come $u+10$ quando u diventa grande[10]. Ma $s = u+v$; quindi $v = s-u$ tende a 10 quando u tende all'infinito — che fisicamente vuol dire che quando l'oggetto scappa all'infinito l'immagine si avvicina quanto vogliamo al fuoco della lente.

Dunque s tende all'infinito sia quando u tende a $+\infty$ sia quando u tende a 10^+. Ma allora la funzione s non può essere monotona nella semiretta $(10, +\infty)$; inoltre, scappando all'infinito a entrambi gli estremi di $(10, +\infty)$, sembra molto ragionevole supporre che abbia almeno un punto di minimo all'interno di questo intervallo. Questo punto di minimo è chiaramente interessante, in quanto ci permette di trovare la distanza minima fra l'oggetto e l'immagine; vedremo nel prossimo capitolo come fare per (dimostrare che esiste e) trovare questo punto. La Figura 4.12 contiene il grafico della funzione s assieme all'asintoto verticale e all'asintoto obliquo[11].

4.7 Funzioni esponenziali

4.8 Funzioni logaritmiche

4.9 Tecniche di interpolazione

4.10 Funzioni trigonometriche

[10] E infatti il grafico di s si avvicina sempre più alla retta grafico di $u + 10$; si dice che questa retta è un *asintoto obliquo* per il grafico di s.

[11] Sì, il grafico di s è un'iperbole non equilatera.

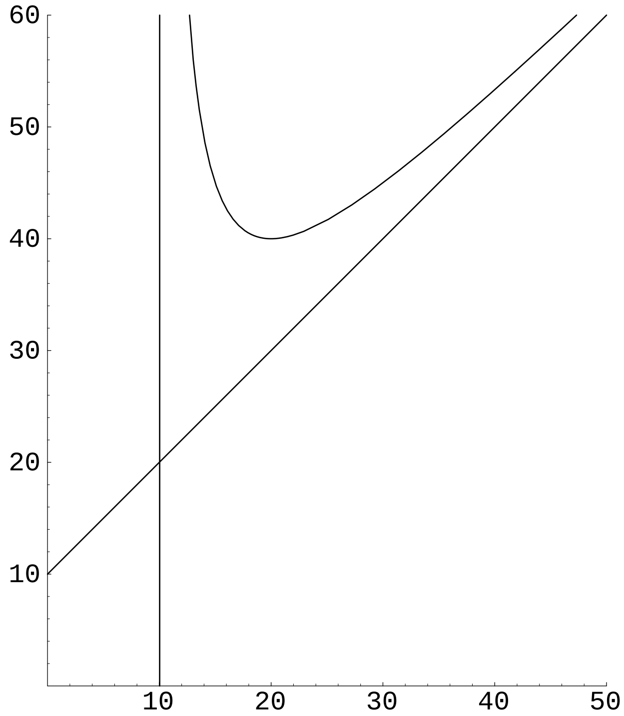

Figura 4.12 $s(u) = u^2/(u-10)$.